AD_____

Award Number: DAMD17-95-2-5017

TITLE: Conformationally Restricted Synthetic AIDS Vaccine

PRINCIPAL INVESTIGATOR: Arnold C. Satterthwait, Ph.D.

CONTRACTING ORGANIZATION: The Burnham Institute
La Jolla, California 92037

REPORT DATE: February 2001

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 2001	Final (1 Jul 95 - 31 Jan 01)

4. TITLE AND SUBTITLE
Conformationally Restricted Synthetic AIDS Vaccine

5. FUNDING NUMBERS
DAMD17-95-2-5017

6. AUTHOR(S)
Arnold C. Satterthwait, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Burnham Institute
La Jolla, California 92037

E-MAIL: asat@burnham.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

It was proposed to develop a structure-based approach for synthetic vaccine development designed to recapitulate the activities of potent neutralizing monoclonal antibodies (MAbs) using constrained peptides as immunogens. We focused on two antibodies, a MAb 58.2 that binds a sequential V3 epitope and IgG b12 that binds a discontinuous epitope, both on HIV-1 gp120. We validated the structure-based approach with a V3 mimetic using MAb 58.2 and demonstrated the enormous effect that peptide conformation can have on affinity and immunogenicity. We were unable to mimic the more challenging discontinuous IgG b12 epitope despite considerable effort.

The enormous affinity enhancements we observed for the constrained V3 peptide and dearth of potent HIV-1 neutralizing antibodies, particularly those specific for sequential epitopes, led us to an important discovery. It suggested that new antibodies might be identified with constrained peptides that go undetected with peptides. It was found that several helix-stabilized gp120 peptides detect new antibodies that go undetected with linear peptides implying a general phenomenon. Constrained peptides, tailored to mimic short, sequential regions of proposed neutralization sites might be used to widen the selection of template antibodies best suited for synthetic vaccine development.

14. SUBJECT TERMS
HIV-1, IgG b12, MAb 58.2, synthetic vaccine, vaccine, peptides
cyclic peptides, V3 loop, gp120, gp41

15. NUMBER OF PAGES
67

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

TABLE OF CONTENTS

INTRODUCTION

We originally hypothesized that HIV-1 had developed a capability for evading immune responses that could be overcome by focusing immune responses to specific targets essential for the virus and defined in part by neutralizing antibodies and T-cell responses. To test this hypothesis, we proposed to identify synthetic mimetics corresponding to neutralizing epitopes recognized by broadly neutralizing antibodies as well as seek new antibodies using mimetics. We proposed to initially focus on two antibodies, MAb 58.2 that binds a sequential epitope on the V3 loop and IgGb12 that binds a discontinuous epitope which overlaps with the CD4 binding site. Both epitopes reside on the HIV-1 envelope protein gp120.

Technical Objectives

Overall Objective (taken from original proposal) : The overall objective is to (1) identify the fine structure specificity and conformational requirements of neutralizing epitopes. (2) mimic these epitopes with synthetic peptides and/or constrained peptides, (3) combine them with T-cell epitopes in MAPS and (4) use them alone or in combination to formulate an AIDS vaccine.

Specific Objectives (original objectives in italic are followed by a summary of the results with reference to detailed descriptions):

1. *Identify conformational epitopes recognized by HIV neutralizing antibodies using loop libraries expressed on filamentous phage.*

a. Ten phage-antibody libraries were used as a source of constrained loop libraries with over a billion independent clones for screening against a high-affinity version of IgGb12, Fab 3B3. Although several clones bound the antibody and were sequenced, there was only passing homology with gp120 sequences (See Army Report 1996, Ref. 1a for details).

b. A reverse screen, using a cyclic C3 peptide, [JKQSSGGDPEIVGZ]C, which encompassed an epitope critical for CD4 and IgGb12 binding was found to select for the b13 family of CD4-binding site antibodies from Fab-phage libraries derived from HIV-1 infected individuals (See Army Report 1996, Ref. 1a for details). However, affinity must be weak since we could not establish direct binding of the cyclic C3 peptide to isolated Fab b13. The cyclic C3 peptide did not select for IgGb12 (See Army Report 1996, Ref. 1a for details). The cyclic C3 peptide was tested as an immunogen (See #3a).

c. A cyclic disulfide loop that binds IgG b12 was identified by our collaborator, Dr. Jamie Scott, using phage-peptide libraries. This result was confirmed in our lab by showing that the corresponding synthetic disulfide loop, CJS2, bound weakly to IgGb12 (See Army Report 1997, Ref. 2a for details). However, there was no correspondence between the amino acid sequence of the disulfide loop and epitopes on gp120 required for IgGb12 affinity.

d. Focused peptide libraries based on shuffled gp120 residues important for IgGb12 binding were synthesized. These libraries were screened with IgGb12. Several peptides, rich in tryptophan were identified that blocked gp120 binding to sCD4 and IgGb12 at 10-100 μM (See Army Report 1996, Ref. 1c for details). There was no apparent relationship between these peptides and a gp120 sequence.

To summarize, when IgGb12 was screened with several different types of large peptide libraries, different low affinity peptides were selected with at best only passing homology with

gp120. This suggests that either peptides are being selected by irrelevant binding cavities on IgGb12 and/or tolerance at the binding site for amino acid diversity which complicated analysis and further development. However, a definitive interpretation will require X-ray crystallography of Fab b12-peptide complexes. The cyclic disulfide loop, CJS2, has been submitted to Wilson's lab for crystallization with the Fab fragment of IgG b12 (See #4b).

2. *Develop predictions for secondary and tertiary structures of neutralizing epitopes on HIV-1 gp120 and gp41 proteins using computer assisted analyses of primary amino acid sequences.*

We collaborated with Skolnick's lab to predict structures for the V1/V2 (See Army Report 1996, Ref. 1b for details) and V3 loops (see original proposal) providing structural data that complements the missing parts of the core crystal structure for HIV-1 gp120 bound by sCD4 and a neutralizing Fab 17b that has been reported by Hendrickson's group [14,15]. Crystal structures for HIV-1 core gp41 have been reported by several groups.

3. *Conformationally restrict selected peptides from gp120 and gp41 to loops identified with libraries and to predicted structures using covalent hydrogen bond mimics and machine-assited multiple cyclic peptide synthesis (MCPS).*

Several extensive studies were carried out to identify constrained peptide mimetics of neutralizing epitopes. The following constrained peptide mimetics were synthesized.

a. A cyclic C3 peptide, [JKQSSGGDPEIVGZ]C, that encompassed an epitope, DPE, recognized by IgGb12 on gp120 was synthesized (See Army Report 1996 and Ref. 1d for details). Although this cyclic peptide did not bind IgGb12, it selected for the b13 class of CD4-binding site antibodies from Fab-phage libraries (#1b). It was prepared and tested as an immunogen (See #7,#8).

b. A cyclic disulfide loop peptide, CJS-2, selected from a phage-peptide library by IgGb12 (#1c) was synthesized and shown to bind weakly to IgGb12. Binding was specific and dependent on cyclization (See Army Report 1997, Ref. 2a for details). Although intriguing, it showed no correspondence with a gp120 sequence. The cyclic peptide was submitted to Wilson's group for X-ray crystallography with the Fab fragment of IgG b12 (#4b).

c. Linear and cyclic C3 peptides encompassing the DPE epitope were linked through peptides of various lengths to ADSRDNWR from C5 which is also likely bound by IgGb12. However, these peptides did not bind IgGb12 (See Army Report 1997, Ref. 2a for details).

d. A small quantity of peptide corresponding to the V1/V2 domain was synthesized (See Army Report 1997, Ref. 2b for details).

e. A cyclic V3 peptide that mimics the conformational epitope bound by MAb 58.2 was optimized for affinity (See Army Report 1998, Ref. 3 as well as Ref. 5 for details). Several cyclic V3 peptides have been submitted to the Wilson lab for crystallography (#4a).

f. A series of gp120 peptides were stabilized as α-helices using the hydrazone link. Parren's lab at The Scripps Research Institute has shown that several of these stabilized α-helices identify antibodies in HIV-1 infected human antisera that are poorly recognized if at all with corresponding linear peptides (See this Report, pp 9-19).

4. *Collaborate on the X-ray crystal structure determination of constrained peptides bound to neutralizing antibodies.*

a. A crystal structure for a long linear and two hydrazone-linked, cyclic V3 peptides bound to MAb 58.2 have been reported (Refs. 4, 5). A high affinity cyclic V3 peptide was submitted for crystallography with MAb 58.2.

b. A cyclic disulfide loop that binds IgG b12 (#1c, #3b) has been submitted to the Wilson lab for crystallization.

5. *Validate constrained peptides as antigens in binding assays with receptors and/or neutralizing antibodies to gp120 and gp41.*

See # 3.

6. *Determine the solution structure of promising constrained peptide vaccines by nuclear magnetic resonance (NMR) spectroscopy when warranted for rational vaccine design.*

NMR spectroscopy was used to compare solution structures of a linear and cyclic V3 peptide that showed low and high affinity for MAb 58.2 respectively (See Army Report 1998, Ref. 3 and also Ref. 5 for details). It was found by NMR that the cyclic peptide stabilized solution conformations that closely resembled the MAb 58.2 bound conformer that had been determined by X-ray crystallography [4,5]. This demonstrates that the enhanced affinity of a constrained peptide for an antibody reflects mimicry of the bound epitope. Enhanced affinity thus provides a simple screen for identifying constrained peptides mimetics of neutralization sites.

7. *Combine constrained peptides with T-cell epitopes in multiple antigen presentation systems (MAPS).*

Two corresponding pairs of linear and constrained peptides were linked to a core MAPS containing a "promiscuous" tetanus toxoid T-cell epitope. These large peptides (>130 amino acids) were prepared in nearly pure form as judged by mass spectrometry and generated high titers of antibodies in mice and rabbits.

a. Linear and corresponding cyclic C3 peptide-MAPS (#1b,3a)(See Army Report, Ref. 1d for details).

b. Linear and corresponding cyclic V3 peptide-MAPS (#3d,4a,6)(See Army Report, Ref. 3 and Ref. 5 for details).

8. *Evaluate immungenicity of MAPS in animals and neutralization titers with lab strains and non-passaged primary isolates.*

a. The cyclic C3 peptide-MAPS (#7a) generated polyclonal antibodies in mice that bound gp120 IIIB while antibodies generated to the linear C3-MAPs did not bind gp120. The result was confirmed by BIACORE (See Army Report 1996, Ref. 1d for details). However, subsequent tests in mice and rabbits did not generate gp120 IIIB or MN binding antibodies.

b. Rabbit polyclonal antibodies generated to the cyclic V3-peptide MAPS but not the linear V3 peptide-MAPS bound gp120 (MN) and neutralized HIV-1 (MN). However, these antibodies did not neutralize a primary isolate, JR-CSF (See Army Report 1998 and Ref. 5 for details).

Identifying Antibodies to Conformational Epitopes on HIV-1 gp120 with Alpha-Helical Peptides

Edelmira Cabezas, Paul W. H. I. Parren and Arnold C. Satterthwait

Antisera are often screened with peptides in attempts to identify native-protein epitopes for vaccine design. However, most native-protein antibodies do not bind peptides and those that do often show preferences for denatured protein. We have identified antibodies in antisera from HIV-1 infected individuals using peptides stabilized as alpha-helices that are poorly recognized by linear peptides. This data indicates that a novel class of conformationally sensitive antibodies can be identified with constrained peptides. In the present case, these include antibodies to an alpha-helix on the "immunologically silent face" of gp120 and a conserved CD4-induced CCR5 binding site on gp120.

The structure-based approach to synthetic vaccines relies on using potent neutralizing monoclonal antibodies as templates for developing constrained peptide mimetics [5]. While this approach can be effective, there are few neutralizing monoclonal antibodies that neutralize primary isolates. On the other hand, antisera from a subset of long-term nonprogressors show broadly neutralizing activities [6] implying that additional neutralizing antibodies remain to be isolated and characterized .

Antibodies generated by native proteins often show conformational requirements for discontinuous epitopes [7]. Discontinuous epitopes present a considerable challenge for peptide mimicry. For this reason, researchers have searched for peptide-reactive antibodies since in some cases the peptides they bind can be utilized to recapitulate neutralizing activities [8]. However, this approach has often led to disappointment. It has also been argued that peptide-reactive antibodies reflect antibodies generated by denatured protein and are irrelevant [7]. This viewpoint is in fact strongly supported by peptide-reactive monoclonal antibodies (MAbs) generated by HIV-1 gp120 which mostly show a preference for denatured gp120 compared with native gp120 [9].

Two of our earlier studies showed that constrained peptides, designed to mimic protein loops, could detect antibodies generated by epidermal-growth-factor-like proteins that went undetected with linear peptides [10,11]. This implied that the affinities of native-protein antibodies could be very sensitive to peptide conformation. We then showed that the affinity of a V3 peptide for the HIV-1(MN) neutralizing antibody MAb 58.2 could be

enhanced >1,000-fold by constraining it [5]. It was further shown in structural studies that the constrained V3 peptide stabilized conformation(s) that reflected the MAb binding cavity [5]. The large affinity enhancements of constrained peptides for native-protein antibodies suggested that this effect might be widespread. If so, then new antibodies might be detected in HIV-1 neutralizing antisera using constrained peptides corresponding to envelope protein substructures. Antibodies that show conformational sensitivity likely reflect native protein.and could be used as templates for structure-based approaches. To determine whether conformationally sensitive antibodies, detectectable with constrained peptides might be a common feature of HIV-1 antisera, we screened antisera from long-term nonprogressors with a series of peptides from gp120 stabilized as α-helices.

Results

Although α-helices make up about a third of globular proteins including HIV-1 gp120 and are often exposed on protein surfaces, we are unaware of any rigorous studies establishing that native-protein binding antibodies can bind short peptides from α-helical regions. This runs counter to the general notion that the whole surface of a protein should be antigenic [7]. Since the hydrazone link can be used to stabilize peptides as α-helices, we inserted it into peptides corresponding to every predicted and observed α-helix in gp120 for testing as antigens. If the whole surface was antigenic, then perhaps new activities might be detected.

The hydrazone link is ideal for stabilizing peptides as α-helices for screening purposes [12] since it can be used to replace an i, i + 4 hydrogen bond at the N-termini of peptides to form an α-helix nucleation site (NucSite)[12]. NucSites overcome the energetically most significant barrier to helix stabilization and propagate peptides as α-helices without alterations of side chains leaving them free for binding to antibodies. We synthesized a series of overlapping pairs of corresponding linear and α-helix nucleated HIV-1_{MN} peptides (A1-A6 and V3, 10-12 amino acids) (Table 1) based on predicted gp120 structures [13]. The MN peptides were inserted between acetyl-GLAGA or the corresponding α-helix nucleation site (NucSite-Ala, [**J**LA**Z**]A) and K(Biotin)-NH_2. Since sequence determines the degree to which a peptide can be propagated as a helix, we selected overlapping sequences on the basis of helix propensities and occasionally replaced MN amino acids with Ala which stabilizes the helical conformation. **J** and **Z** are linkers used to form the hydrazone link during solid phase sysnthesis [12]. The C-terminal biotin was used to adhere peptides to neutrAvidin coated microtiter wells for antibody binding.

The α-helix stabilized peptides include those that overlap corresponding helices in a crystal structure of a gp120 core complexed with two-domain sCD4 and Fab 17b [14,15]. The core gp120 structure was stripped of N and C-terminal sequences and the V1/V2/V3 loops while retaining the V1/V2 stem. Fab 17b binds a CD4-inducible epitope which overlaps the CCR5 chemokine receptor binding site and blocks binding of CD4 bound gp120 to CCR5 which is an obligatory step in HIV-1 infection.

As summarized in Table 2, the helix-stabilized peptides encompass sequences from α-helices from the inner domain (α1) of the complexed core, from exposed V2 (α2) and V3 (V3) regions that were absent from the gp120 core, from the "immunologically silent face" of gp120 (α3), present on the core, from a region (α4) that adapts a strand-β-hairpin structure on the core, from a second inner domain α-helix (α5) present in the core and from the C-terminal end that is truncated from the core. Two short α-helices (≤ 5 amino acids) present in the core structure were not predicted and deemed to short for stabilization with a NucSite.

The α4-1 peptide (Table 1) encompasses amino acids that overlap epitopes for CD4bs antibodies (K421,W427), sCD4 (I424), MAb 17b (R419,I420,K421,Q422,I423) used to cocrystallize core gp120 and the CCR5 binding site (R419,I420,K421,Q422). Although this region is non-helical in the complexed gp120 core, it is central to a CD4-induced epitope that is bound by MAb 17b and CCR5.

Two human antisera were titered against the peptides and $rgp120_{MN}$. These antisera, FDA-2 and RW1, showed broadly neutralizing activity against primary isolates from several Clades [16]. The FDA2 antisera showed higher titer against rgp120 which are reflected by titers to the peptides. Taking into account the reduced titers for the RW1 antisera, the profiles were essentially identical. The titration data for the FDA-2 antiserum for linear and helix-stabilized gp120 peptides are plotted in Figure 1.

The stabilized α-helical peptides identify new antibody activities that are barely observed if at all with the linear peptides (Fig. 1, summarized in Table 1). These include antibodies to an α-helix on the "immunologically silent" face of gp120 (α2 on core gp120) and to a region within the CD4i epitope (predicted α4 helix). On the other hand, no reactions were observed with α-helical peptides corresponding to the two inner domain α-helices (α1, α5 on core gp120) present in the core structure nor to the V2 peptides (predicted α2). A weak reaction was observed with the V3 helical peptide. Antibodies were detected with both the linear and helical C-terminal peptides (predicted α6).

The enhanced reactions with α-helical peptides imply that the bound antibodies were generated to α-helices. They demonstrate that constraining peptides to α-helices can significantly enhance their activities (affinities) for antibodies generated to α-helices. Most

helices in proteins are short with a median length of 10 residues (88% < 15 residues) and devoid of structure in water even at low temperature. Since the affinity of peptides to antibodies is proportional to the degree by which they adapt a complementary structure in solution [5], the results imply that many reactions of antibodies generated against α-helices will be missed by screening with short peptides. These results complement preliminary results with constrained peptide loops which showed enhanced activities with antibodies generated by *P. falciparum* EGF-like proteins. The enhanced affinity of constrained peptides for antibodies provide a means for distinguishing native-protein antibodies from denatured protein antibodies.

Most but not all of the data can be accounted for in terms of what is known about gp120 from the core crystal structure [14,15] and other data. No antibodies were detected to two inner domain α-helices as would be expected if antibodies were formed against an oligomeric trimer that buried these helices. The inner domain is exposed on monomeric rgp120. Although sequence variation in infective HIV-1 viruses could reduce binding of MN peptides used in this study to antibodies specific for different sequences, the α1 and α5 sequences show very little variation across Clade B isolates and neither antiserum binds these peptides.

On the other hand, antibodies were detected to α3 helix-stabilized peptides. These peptides correspond to an alpha helix (α2) located on the exposed outer domain of the core gp120 crystal structure [14]. Although the gp120 α2 helix is in the "immunologically silent region", the presence of antibodies to this epitope implies that this region is not "silent" but rather antibodies to this region are conformational antibodies that can not be detected with linear peptides. The gp120 α2 helix undergoes considerable variation [14] which agrees with its exposure to antibodies. The titer to the MN constrained peptide could reflect differences in fine structure specificity which would lower it. The "silent" face is adjacent to the "neutralizing" face which binds CD4 and CCR5 [14]. Although it's location implies that the "silent" face is nonneutralizing, the potent neutralizing MAb 2G12 shows specificity for glycosylated Asn residues on this face [17]. Helix-stabilized peptides corresponding to α2 on core gp120 might be used for isolating antibodies to the silent face for determination of whether it is a neutralizing face.

These former results which can be accounted for in terms of the core gp120 crystal structure contrast to the binding of the helical A4-1 and A4-2 α4 peptides. The cognate sequences in the core gp120 form a non-helical structure that is bound by Fab 17b [14,15]. The A4-1 sequence encompasses amino acids that correspond to β-strands, β19 (RIK) and β20 (QIINM) that forms at an angle to β19 and a loop (NMWQ). These regions encompass a core region within the CD4i MAb17b and CCR5 binding site epitopes (K/R419-I423).

The A4-2 peptide corresponds to the apex of a β-hairpin formed from β20 (INM), a loop (NMWQK) and β21 (VGKA). CD4 binds N425-V430 at the apex of the β-hairpin, which is represented in A4-1 (NMWQ) and A4-2 (NMWQEV); the MN sequence substitutes E for K present in the core which is from the BH10 virus. There is more potential for variability in the A4-2 sequence which also includes a G/A substitution than in the A4-1 sequence which could account for its lower titer.

The observation that antibodies bind α-helical A4 peptides implies that the adjacent CD4/CCR5 binding regions can undergo a conformational transition, possibly between an α-helix on native gp120 and CD4-bound gp120. This agrees with prior observations that suggest a conformational transition takes place [14,18]. Studies show that when sCD4 binds gp120, epitopes on the V3 loop and the chemokine receptor binding site on gp120 are exposed (CD4i epitopes) to binding by CD4i antibodies and CCR5 indicating a CD4 induced conformational change [18]. The changes could involve the repositioning of loops and/or more elaborate structural changes. The observation of antibodies specific for α4 α-helical peptides is consistent with a transition from a C4 α-helical structure to the more complex core structure that is observed to be bound by sCD4. However, if a structural reorganization occurs, it may be localized as suggested by the presence of antibodies to the gp120 α2 helix which is present on the core structure.

Discussion

This study demonstrates that new antibodies can be detected in HIV-1 antiserum with α-helix nucleated gp120 peptides. Earlier, we reported that cyclic peptide mimetics of protein loops identified antibodies in antisera generated to native EGF-like proteins [10,11]. The observations that constrained peptides nucleated as α-helices and loops can identify antibodies in antiserum from different sources implies that this novel class of antibodies is widespread. The detection of these conformationally sensitive antibodies bring perspective to peptide-reactive antibodies and has important implications for synthetic vaccine research.

The new antibodies show large affinity enhancements for constrained peptides compared with linear peptides. It is likely that they reflect antibodies generated to native protein. These large affinity enhancements are mirrored by the affinity enhancements of constrained V3 peptides for MAb 58.2 [5]. In this latter case, NMR spectroscopy [5] and X-ray crystallography [19] was used to demonstrate a conformational basis for this effect. Constraining a peptide prior to binding reduces the energy cost and enhances affinity [5].

Linear peptides are often used to detect antibodies and analyze immune responses to pathogens as a prelude to vaccine studies [7,9]. However, the origins and significance of

these activities have been hotly debated [7]. It have often been observed that one or a few overlapping peptides based on a protein sequence are antigenic while many others show no activity. This has led to the concept of immunodominance i.e. that certain regions of proteins are more immunogenic than others. A completely different position regarding peptide-reactive antibodies has been taken by others [7]. This position is that peptide-reactive antibodies mostly reflect denatured protein while most antibodies generated by native protein bind discontinuous epitopes [7]. The large effect of conformation on peptide affinity for serum antibodies shed light on these issues.

It has been noted that peptides corresponding to "immunodominant" epitopes that are immunogenic frequently adapt β-turn conformations at low temperature in water [20]. Our data implies that any tendency of a peptide towards stabilization of a protein structure would significantly enhance affinity for native-protein antibodies. Presumbly, the detection of a limited number of peptides by some antisera reflect the tendency of a small number of peptides to partially stabilize protein like conformers. The detection of serum antibodies with peptides that stabilize β-turns thus likely reflects immunodetectability rather than "immunodominance". If so, then peptides stabilized as loops and α-helices should cast a wider net as we have observed.

The discovery that constrained peptides bind antibodies with enhanced affinity, identifying new activities could have important ramifications for synthetic vaccine research in general and HIV-1 vaccine research in particular. First, constrained peptides identify new antibodies suitable for use as templates in structure-based synthetic vaccine strategies. Second, enhanced affinity provides a means for distinguishing antibodies generated by native protein from those generated by denatured protein. Most peptide-reactive MAbs generated to gp120 show preferences for denature gp120 [9]. Third, affinity enhancement provides a simple measure for optimizing the structures of peptide mimetics as demonstrated in the MAb 58.2 study [5] which in turn are demonstrably far more immunogenic than corresponding linear peptides.

A major challenge for an HIV-1 vaccine will be the genetic variability of the virus. It may be of critical importance to develop vaccines that can focus immune responses on epitopes least subject to mutation. Constrained peptides provide a means for focusing antibodies on defined epitopes that may be difficult to achieve by other means. A major problem for synthetic vaccine researchers has been a lack of template antibodies for use in optimizing peptide immunogens. The discovery that a whole new class of conformationally sensitive antibodies can be detected with constrained peptides might be utilized to alleviate this need.

Materials and Methods

Synthesis. Linear and helical peptides were synthesized by the solid-phase method, following the Fmoc strategy. Side-chains were protected using Fmoc-Arg(Pbf), Fmoc-Asp(OBut), Fmoc-Glu(OBut), Fmoc-His(Trt), Fmoc-Lys(Boc), Fmoc-Ser(But), Fmoc-Thr(But), Fmoc-Trp(Boc) and Fmoc-Tyr(But). Peptides were analyzed and purified with Gilson analytical and preparative HPLC systems using linear gradients formed from solvent systems A: 0.1%TFA/H_2O (v/v) and B: 0.1%TFA/AN (v/v).

Linear peptides were synthesized with an ACT-350 multiple peptide synthesizer (Advanced Chem Tech) on Fmoc-Lys(ε-d-Biotin-LC)-Rink amide-MBHA resin (0.34 meq/g) using standard Fmoc synthesis with DIC/HOBt coupling. The N-termini of peptides were acetylated on the resin with 15% acetic anhydride in DMF. The acetylated peptides were cleaved with Reagent K *(1)*, precipitated with diethyl ether and purified using a preparative RP C-18 column (Cosmosil 5C18-AR, 2 x 25 cm) eluting at 8 mL/ min with a gradient of 0-20% AN over 5 min, followed by 20-60% AN over 25 min. Peptides were confirmed by electrospray mass spectroscopy.

Helical peptides were synthesized by using a hydrazone link which was inserted into peptides during solid-phase synthesis using Fmoc-Lys(e-d-Biotin-LC)-Rink amide-MBHA resin (0.34 meq/g) according to Scheme 1 using previously described procedures [12] These syntheses required the prior preparation of J and FmocZ(Act) *(2)*. N-α-Fmoc-Alanine chloride was prepared according to Carpino et al. [21] and shown to be > 98% acyl chloride by the methanol test. Initially, the peptides were synthesized on the synthesizer up to the residue that precedes Z as described above for linear peptides. The peptide-resin was then transferred to mesh packets to continue the synthesis of JLAZ using a previously described protocol [12]. The peptides were cyclized on the resin using HCl (2 meq of HCl/meq of peptide) in 20% TFE/DCM [12]. Cyclic peptides were cleaved from the resin and deprotected with 5% H_2O/95% TFA and precipitated with diethyl ether. The cyclic peptides were purified by HPLC using a preparative RP C-18 column (Cosmosil, 5C18-AR) eluting at 8 mL/min with a 0-20% AN gradient over 5 min followed by a 20-50 or 60% AN over 25 min. The purified peptides were confirmed by electrospray mass spectroscopy.

ELISA assay. Linear and helix peptides (100 ng/100 μL PBS/well) were added to 96-well neutrAvidin coated microtiter plates (Pierce) and incubated overnight at 4°C, the plates were then washed four times with distilled water. FDA-2 and RW1 human antiserum from a long-term nonprogressor [16] were serially diluted two-fold with 1% BSA/0.02% Tween 20/PBS into the wells and incubated for 1 hour at 37°C. After washing 4 times with 0.05%

Tween 20/PBS, bound antibody was determined using alkaline phosphatase-conjugated goat antihuman $F(ab)_2$ (Pierce) (50 µL/well of 1/500 dilution of $F(ab)_2$ in 1% BSA/0.02% Tween 20/PBS. After incubation at 37°C for one hour and washing four times with 0.05% Tween 20/PBS, 50 µL of developing solution (1 mg pNPP/ mL developing buffer) was added. Optical density was determined at 405 nm after 15 or 40 min with a Molecular Devices SpectraMAX.

Seq. Range	AA sequence	Pep. #	Ab	
91-114	**VNVTENFNMWKNNMVEQMHEDIISLWDQS**		**α1**	
	Ac-GLAGATENFNMWKNNKB	1581	(-)	
	JLAZATENFNMWKNNKB	1527	(-)	(A1-1)
	Ac-GLAGANMVEQMHEDIKB	1582	(-)	
	JLAZANMVEQMHEDIKB	1528	(-)	(A1-2)
	Ac-GLAGAEDIISLWDQSKB	1583	(-)	
	JLAZAEDIISLWDQSKB	1529	(-)	(A1-3)
170-181	**IRDKMQKEYALL**		**α2**	
	Ac-GLAGARDKMQKEYALKB	1584	(-)	
	JLAZARDKMQKEYALKB	1530	(-)	(A2)
339-357	**RAKWNDTLRQIVSKLKEQF**		**α3**	
	Ac-GLAGARAKWNDTLRQIKB	1585	(+)	
	JLAZARAKWNDTLRQIKB	1532	(++)	(A3-1)
	Ac-GLAGARQIVSKLKEQFKB	1586	(-)	
	JLAZARQIVSKLKEQFKB	1533	(+)	(A3-2)
418-433	**CKIKQIINMWQEVGKA**		**α4**	
	Ac-GLAGAKIKQIINMWQKB	1587	(+)	
	JLAZAKIKQIINMWQKB	1534	(+++)	(A4-1)
	Ac-GLAGAINMWQEVAKAKB	1588	(-)	
	JLAZAINMWQEVAKAKB	1535	(++)	(A4-2)
478-487	**RDNWRSELYK**		**α5**	
	Ac-GLAGARDNWRSELYKYKB	1589	(-)	
	JLAZARDNWRSELYKYKB	1536	(-)	(A5)
502-513	**KAKRRVVQREKR**		**α6**	
	Ac-GLAGAKAKRRVVQREKRB	1590	(++)	
	JLAZAKAKRRVVQREKRB	1537	(++)	(A6)
328-335	**GTIRQAHC**		**V3**	
	Ac-GLAGARQAHANISRAKB	1591	(-)	
	JLAZARQAHANISRAKB	1538	(+)	(V3)

Table 1. Predicted alpha-helices for gp120 (MN). Ac is acetyl; B is Lys (linker-d-biotin). JLAZ is a cyclic peptide that nucleates alpha-helix formation [12]. Peptides are numbered for reference to the titer data in Fig. 1. Alpha-helix nucleated peptides are also lettered (A1-1, etc.) for reference to Table 2. Antibody titers are summarized under the Ab column; (-), no titer; (+) questionable titer; (++,+++), increasing titer.

Comparison of Predicted and gp120 crystal structure α-helices			
		complexed core gp120 crystal structure	
Predicted α-helices	**Peptide**	**Observed α-Helices**	**Structure**
α1	A1-1	α1	irregular
	A1-2		buried helix
	A1-3		buried helix
α2	A2	from truncated V2	
α3	A3-1	α2	helix (silent face)
	A3-2		helix (silent face)
		α3	short helix (5 residues)
		α4	short helix (4 residues)
α4	A4-1		β-strands, 90° bend
	A4-2		apex of β-hairpin
α5	A5	α5	buried helix
α6	A6		truncated, frayed end
V3	V3		truncated

Table 2. Comparison of predicted and observed a-helices in gp120. The α-helices were observed in core gp120 complexed by sCD4 and Fab 17b that binds a CD4 induced epitope [14,15].

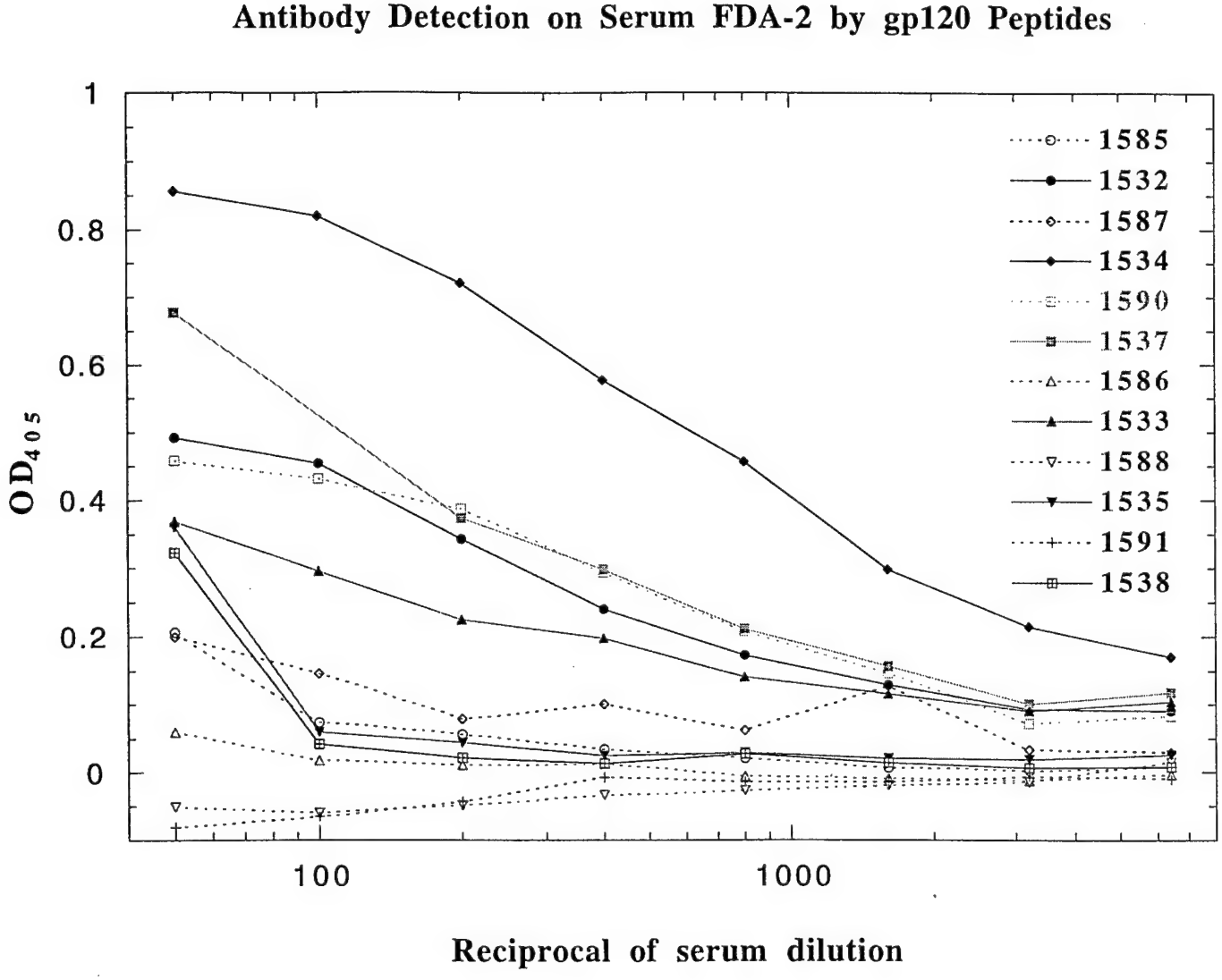


Figure 1. Titration of FDA-2 antiserum with corresponding pairs of linear (dashed lines) and helix stabilized (solid lines) gp120 peptides. Corresponding pairs of peptides share the same symbol. Peptides sequences are identified in Table 1.

KEY RESEARCH ACCOMPLISHMENTS

- Demonstrated that the hydrogen bond mimic approach to constraining peptides is applicable to a wide range of protein sequences and different conformations including α-helices and loops which make up about 50% of native protein surfaces.

- Demonstrated that peptides stabilized as α-helices and loops show enormous affinity enhancements for antibodies.

- Validated the structure-based approach to synthetic vaccine development using the V3 MAb 58.2 in the most complete study of its kind to be published to date

- Demonstrated that new, conformationally sensitive HIV-1 antibodies that are likely better suited for use as templates in synthetic vaccine development can be identified with constrained peptides

REPFORTABLE OUTCOMES

1. Presentation at the 15th American Peptide Symposium, Nashville, TN (1997): Conformational Mapping of Neutralizing Epitopes with Peptide Mimetics.

2. Presentation at the 5th Chinese Peptide Symposium, Lanzhou, China (1998): Polyclonal immune response to a constrained V3 peptide mimics the conformational preference and neutralizing activity of a potent HIV-1 neutralizing MAb.

3. Cabezas, E., and Satterthwait, A.C. (1998) The NMR structure of a V3 loop peptide that binds tightly to a monoclonal antibody that potently neutralizes HIV-1 . In: Peptides: Chemistry, Structure and Biology. Proceedings of the 15th American Peptide Symposium. J. P. Tam and T. P. Kaumaya (eds.), ESCOM Leiden, The Netherlands, pp 441-442.

4. Stanfield, R.L., E. Cabezas, A. C. Satterthwait, E.A. Stura, and I.A. Wilson, (1999) Dual conformations for the HIV-1 gp120 V3 loop as seen in bound complexes with different neutralizing Fabs, *Structure*, 7: 131-142.

5. Cabezas, E., Wang, M., Parren, P. W. H. I., Stanfield, R.L., and Satterthwait, A. C., (2000) A structure-based approach to a synthetic vaccine for HIV-1, *Biochemistry,* 39: 14377-14391.

6. Dovalsantos, E. and Satterthwait, A.C., .Dovalsantos, E. and Satterthwait, A.C., Synthesis of peptides incorporating helix mimetics and inducers, In "Synthesis of Peptidomimetics" Houben-Weyl Vol. E22b for Methoden der Organic Chemische, A. Felix, L. Moroder and C. Toniolo (eds), Georg Thieme Verlag, Stuttgart, accepted

7. Cabezas, E., Calvo, J.C. and Satterthwait, A. C. (2001) Constrained peptides detect a novel class of antibodies. Abstract submitted to 17th American Peptide Symposium

8. Cabezas, E. and Satterthwait, A.C. (2001) HIV-1 glycoprotein-120 peptides stabilized as alpha-helices identify a novel class of antibodies in HIV-1 infected human antisera. Abstract submitted to 17th American Peptide Symposium

9. Arnold C. Satterthwait was appointed an Associate Professor at The Burnham Institute in 1999.

10. Arnold C. Satterthwait submitted two proposals to the National Institutes of Health for continuation of HIV-1 vaccine research.

CONCLUSIONS

Our work with MAb 58.2 led to a full development and evaluation of the structure-based approach [5]. It demonstrated the enormous effect that constrained a peptide can have on the affinity of a peptide for an antibody and its immunogenicity. However, while antibodies generated to the V3 mimetic were able to potently neutralize a lab-adapted HIV-1 (MN) isolate they were unable to neutralize a primary isolate. We were unable to mimic the more challenging discontinuous epitope bound by IgGb12 implying that the structure-based approach should be limited to sequential epitopes.

As it stands, there are few peptide-reactive neutralizing antibodies to any pathogen, including primay HIV-1 isolates. This could reflect in part, observations that most peptide-reactive antibodies are generated by denatured protein [7,9]. However, the enormous affinity enhancement of constrained peptides for MAb 58.2 suggested that new native-protein antibodies to sequential epitopes might be identified with constrained peptides. Indeed, we found that HIV-1 gp120 peptides stabilized as α-helices identified new antibodies in HIV-1 infected human antisera that showed weak if any reaction with unconstrained peptides. It is likely these antibodies reflect sequential epitopes on native protein since they display a conformational dependence. This type of antibody is appropriate for optimization of immunogens using the structure-based approach elaborated with MAb 58.2. Consequently, we recommend that more attention be given to identifying and isolating MAbs with conformational dependence for sequential epitopes. The recent structure determinations of gp120 and gp41 provide details for designing constrained peptides corresponding to proposed neutralization sites i.e. CD4 and chemokine receptor binding sites as well as the gp41 inner core coiled coil that can be used for this purpose.

REFERENCES

1. Annual Army Report 1995-1996
a. Carlos Barbas III, Selection of constrained loop libraries for binding to the b12 antibody family and a selection based study of a gp120 C3-region peptide with a antibody library prepared from an HIV-1 infected individual, pp 1-11
b. Wei-Ping Hu, Angel R. Ortiz, Arnold C. Satterthwait and Jeffrey Skolnick, A 3-D structure prediction for the V1/V2 domain of envelope glycoprotein 120 of HIV-1, pp 12-31.
c. Ruoheng Zhang, Ming-zhu Zhang, Yuan Xu, Edelmira Cabezas and Arnold C. Satterthwait, The identification of synthetic peptides that block binding of HIV-1 g120 to CD4 and Fab 12, pp 32-53
d. Arnold C. Satterthwait, Shao-qing Chen and Edelmira Cabezas, The immunogenicity of a constrained peptide from the third constant region of HIV-1 gp120, pp 54-71

2. Annual Army Report 1996-1997
a. Arnold C. Satterthwait, Ming-zhu Zhang, Yuan Xu, Sarah Venturini and Edelmira Cabezas, II. The identification of synthetic peptides that block binding of HIV-1 gp120 to sCD4 and IgG1b12, pp 1-21
b. Ruoheng Zhang, Yuan Xu and Arnold C. Satterthwait, I. Chemical synthesis of the V1/V2 domain from gp120, pp 22-45

3. Annual Army Report 1997-1998
a. A polyclonal immune response to a constrained V3 peptide mimics the conformational preference and neutralizing activity of a potent HIV-1 neutralizing MAb, pp 1-34

4. Stanfield, R.L., Cabezas, E., Satterthwait, A.C., Stura, E.A., Profy, A.T. and Wilson, I.A. (1999) Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing Fabs, *Structure 7,* 131-142. **(In appendix)**

5. Cabezas, E., Wang, M., Parren, P.W.H.I, Stanfield, R. and Satterthwait, A.C., (2000) A structure-based approach to a synthetic vaccine for HIV-1, *Biochemistry 39,* 14377-14391. **(In appendix)**

6. Pilgrim, A. K. et al. (1997) Neutralizing antibody responses to HIV-1 in primary infection and long-term — nonprogressive infection. *J. Infect. Dis. 176,* 924-932.

7. Laver, W.G., Air, G.M., Webster, R.G. and Smith-Gill, S.J. (1990) Epitopes on protein antigens: misconceptions and realities. *Cell 61,* 553-556.

8. Sela, M., and Arnon, R. (1992) *Vaccine 10*, 991-999.

9. Moore, J.P., Sattentau, Q.J., Wyatt, R. and Sodroski, J. (1994) Probing the structure of HIV-1 with a panel of monoclonal antibodies. *J. Virol. 68,* 469-484.

10. Calvo, J.C., Perkins, M., and Satterthwait, A.C. (1994) The identification by chemical synthesis of a conformational epitope on merozoite surface protein 1 of *P. falciparum* malaria. In: Peptides: Chemistry, Structure and Biology. Proceedings of the 13th American Peptide Symposium. R. Hodges and J.A. Smith (eds.) ESCOM Leiden, The Netherlands, pp. 725-726.

11. Satterthwait, A.C., Cabezas, E., Calvo, J.C., Chen, S.Q., Wu, J.X., Wang, P.L., Xie, Y.L., Stura, E.A., and Kaslow, D.C. (1995) A peptide mimetic as antigen and

immunogen. *In:* Peptides: Biology and Chemistry. Proceedings of the Chinese Peptide Symposium 1994. Lu, G. (ed.). ESCOM, Leiden, The Netherlands, pp. 229-233.

12. Cabezas, E. and Satterthwait, A.C., (1999) The hydrogen bond mimic approach: The solid phase synthesis of a peptide stabilized as an α-helix with a hydrazone link. *J. Am. Chem. Soc. 121,* 3862-3875. **(In appendix)**

13. Hansen, J.E. et al. (1996) Prediction of the secondary structure of HIV-1 gp120. *Proteins: Struct. Func. Gen, 25,* 1-11

14. Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., Hendrickson, W. A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature, 393,* 648-659.

15. Kwong, P.D., Wyatt, R., Majeed, S., Robinson, J., Sweet, R. W., Sodroski, J. and Hendrickson, W. A. (2000) Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure, 8,* 1329-1339.

16. Parren, P.W.H.I., Wang, M., Trkola, A., Binley, J.M., Purtscher, M., Katinger, H., Moore, J.P. and Burton, D.R. (1998) *J. Virol.* 72: 10270-10274.

17. Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., Sodroski, J.G. (1998) The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature 393*, 705-711.

18. Rizzuto, C.D., Wyatt, R., Hernandez-Ramos, N., Sun Y., Kwong, P.D., Hendrickson, W.A., Sodroski, J. (1998) A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science 280,* 1949-1953; Rizzuto, C. and Sodroski, J. (2000) Fine definition of a conserved CCR-5 binding region on the HIV-1 gp120. *AIDS Res. Hum. Retro. 16(8),* 741-749.

19. Stanfield, R.L., Cabezas, E., Satterthwait, A.C., Stura, E.A., Profy, A.T. and Wilson, I.A. (1999) Dual Conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing Fabs, *Structure 7,* 131-142. **(In appendix)**

20. Dyson, H. J. and Wright, P. E. (1995) *FASEB J. 9*, 37-42.

21. Carpino, L. A., Cohen, B. J., Stephens,Jr., K. E., Sadat-Aalaee, S. Y., Tien, J. H., and Langridge, D. C. (1986) *J. Am. Chem. Soc. 51*, 3732-3734.

LIST OF PERSONNEL

Arnold C. Satterthwait, Ph.D.
Carlos Barbas III, Ph.D.
Sara Venturini, Ph.D.
Edelmira Cabezas, Ph.D. equiv
Ruoheng Zhang, Ph.D.
Ming-zhu Zhang, Ph.D.
Ellen Dovalsantos, Ph.D.
Weiping Hu, Ph.D.
Yuan Xu, M.Sc.
Gang Liu, Ph.D.
Shiow-Meei Chen, Ph.D.
Julio C. Calvo, M.Sc.

APPENDIX

1. Cabezas, E., Wang, M., Parren, P.W.H.I, Stanfield, R. and Satterthwait, A.C. (2000) A structure-based approach to a synthetic vaccine for HIV-1, *Biochemistry 39,* 14377-14391.

2. Stanfield, R.L., Cabezas, E., Satterthwait, A.C., Stura, E.A., Profy, A.T. and Wilson, I.A. (1999) Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing Fabs, *Structure 7,* 131-142.

3. Cabezas, E. and Satterthwait, A.C. (1999) The hydrogen bond mimic approach: The solid phase synthesis of a peptide stabilized as an α-helix with a hydrazone link. *J. Am. Chem. Soc. 121,* 3862-3875.

Articles

A Structure-Based Approach to a Synthetic Vaccine for HIV-1†

Edelmira Cabezas,‡ Meng Wang,§ Paul W. H. I. Parren,§ Robyn L. Stanfield,|| and Arnold C. Satterthwait*,‡

The Burnham Institute, 10901 Torrey Pines Road, La Jolla, California 92037, and Departments of Immunology and Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received February 16, 2000; Revised Manuscript Received August 14, 2000

ABSTRACT: The generation of neutralizing antibodies by peptide immunization is dependent on achieving conformational compatibility between antibodies and native protein. Consequently, approaches are needed for developing conformational mimics of protein neutralization sites. We replace putative main-chain hydrogen bonds (NH → O=CRNH) with a hydrazone link (N−N=CH−CH_2CH_2) and scan constrained peptides for fit with neutralizing monoclonal antibodies (MAbs). To explore this approach, a V3 MAb 58.2 that potently neutralizes T-cell lab-adapted HIV-1_{MN} was used to identify a cyclic peptide, [JHIGPGR-(Aib)F(D-Ala)GZ]G-NH_2 (loop 5), that binds with >1000-fold higher affinity than the unconstrained peptide. NMR structural studies suggested that loop 5 stabilized β-turns at GPGR and R(Aib)F(D-Ala) in aqueous solvent implying considerable conformational mimicry of a Fab 58.2 bound V3 peptide determined by X-ray crystallography [Stanfield, R. L. et al. (1999) *Structure 142*, 131−142]. Rabbit polyclonal antibodies (PAbs) generated to loop 5 but not to the corresponding uncyclized peptide bound the HIV-1_{MN} envelope glycoprotein, gp120. When individual rabbit antisera were scanned with linear and cyclic peptides, further animal-to-animal differences in antibody populations were characterized. Loop 5 PAbs that most closely mimicked MAb 58.2 neutralized HIV-1_{MN} with similar potency. These results demonstrate the remarkable effect that conformation can have on peptide affinity and immunogenicity and identify an approach that can be used to achieve these results. The implications for synthetic vaccine and HIV-1 vaccine research are discussed.

The development of broadly cross-reactive neutralizing antibodies is an important aim of HIV-1[1] vaccine research (*1*, *2*). However, the HIV-1 envelope glycoproteins, gp120 and gp41, are poorly immunogenic (*1*, *2*). HIV-1 has also undergone an unprecedented degree of mutation (*3*) that complicates vaccine design. Sera from HIV-1 seropositive individuals generally neutralize HIV-1 very poorly. Nevertheless, antisera from a subset of long-term nonprogressors show a marked broadening of HIV-1 neutralizing activities, and a few potent neutralizing monoclonal antibodies have been identified (*2*, *4*, *5*). There is a critical need for approaches to identify and recapitulate these activities.

Antibodies are often identified and their specificities characterized with peptides (*6*). Peptides can in turn be used to generate antibodies which in some cases are neutralizing (*7*). While the availability of peptides favor the approach, it is limited by conformational issues (*8*−*10*). Peptides are conformationally heterogeneous in aqueous solution (*11*), often differing from the structures their cognate sequences adapt in native proteins. This has two important consequences. First, it reduces the affinity of peptides for antibodies generated by native proteins (*8*, *9*, *12*). For a peptide to bind a native protein antibody, it must approximate the native conformation with a cost in free energy that can significantly lower affinity (*12*). Second, the binding pockets of antipeptide antibodies reflect the conformational heterogeneity of peptides (*13*) and are for the most part incompatible with native protein surfaces (*10*, *11*, *13*−*15*).

† This work was supported by the National Institutes of Health (AI37512 to A.C.S. and GM46192 to I.A.W. for R.L.S.) and The Department of Defense, Army (DAMD17-95-2-5017 to A.C.S.).

* To whom correspondence should be addressed: Phone: 858-646-3100, ext. 3658. Fax: 858-646-3195. E-mail: asat@burnham-inst.org.

‡ The Burnham Institute.

§ Department of Immunology, The Scripps Research Institute.

|| Department of Molecular Biology, The Scripps Research Institute.

[1] Abbreviations: TCLA, T-cell lab-adapted; HIV-1, human immunodeficiency virus type 1; MN, HIV-1_{MN}; IIIB, HIV-1_{IIIB}; JR-CSF, HIV-1_{JRCSF}; rgp120, baculovirus-derived recombinant envelope glycoprotein 120_{MN}; p24, HIV-1 protein; MAb, monoclonal antibody; Fab, antigen binding fragment; PAbs, polyclonal antibodies; Aib, α,α-dimethylglycine; NMP, *N*-methylpyrrolidone; DIC, 1,3-diisopropylcarbodiimide; HOBt, 1-hydroxybenzotriazole; DIEA, diisopropylethylamine; DCM, dichloromethane; AN, acetonitrile; TFE, trifluroethanol; MS, mass spectrometry; MAPS, multiple-antigen-peptide system; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline; pNPP, *p*-nitrophenyl phosphate; EC_{50}, effective concentration of peptide that inhibited 50% of the binding of antibody to plate-bound antigen; IC_{90}, concentration (or titer) of antibody that inhibited ≥90% of HIV-1 infectivity; $TCID_{50}$, 50% tissue culture infective dose; NOESY, nuclear Overhauser effect spectroscopy; ROESY, rotating-frame Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; P. E. COSY, phase-edited correlation spectroscopy; TPPI, time-proportional phase incrementation.

10.1021/bi0003691 CCC: $19.00
Published on Web 11/01/2000

Table 1: Linear, Cyclic Peptides and MAPS Prepared for This Study[a]

peptide	sequence
MN	**I1−H2−I3−G4−P5−G6−R7−A8−F9−Y10**
linear 1	Ac-GHIGPGRAFGGG-NH_2
linear 2	Ac-GHIGPGRAFGGGG-NH_2
linear 3	Ac-GHIGP(D-Ala)RAFGGGG-NH_2
linear 4	Ac-GHIGPGR(Aib)FGGGG-NH_2
linear 5	**AcG1−H2−I3−G4−P5−G6−R7−(Aib)8−F9−(D-Ala)10−GGG-NH_2**
linear 5/C	Ac-GHIGPGR(Aib)F(D-Ala)GGC-NH_2
loop 1	[JHIGPGRAFGZ]G-NH_2
loop 2	[JHIGPGRAFGGZ]G-NH_2
loop 3	[JHIGP(D-Ala)RAFGGZ]G-NH_2
loop 4	[JHIGPGR(Aib)FGGZ]G-NH_2
loop 5	**[J−H2−I3-G4-P5-G6-R7-(Aib)8-F9-(D-Ala)10-GZ]G-NH_2**
loop 5/C	[JHIGPGR(Aib)F(D-Ala)GZ]C−NH_2
loop 6	Ac-[CHIGPGR(Aib)FGGC]-NH_2
loop 7	[JHVGPGR(Aib)F(D-Ala)GZ]G-NH_2
MAPS core	(chloro-Ac{TT2}$KKK)_4(KGG)_2$KA-NH_2
linear 5-MAPS	(linear 5/C−Ac-{TT2}$KKK)_4(KGG)_2$KA-NH_2
loop 5-MAPS	(loop 5/C−Ac-{TT2}$KKK)_4(KGG)_2$KA-NH_2

[a] The MN sequence was renumbered; **I1** (our numbering) is I312 from the MN sequence (*3*). The same numbering system was used for the peptides as shown for linear 5 and loop 5. The single letter code is used for standard amino acids. D-Ala is D-alanine; Aib is α,α-dimethylglycine; J and Z are linking residues used to form the hydrazone link which is indicated by brackets; Ac is acetyl; chloro-Ac is chloroacetyl; -NH_2 represents a C-terminal carboxamide and {TT2} is a tetanus toxoid T-helper epitope, QYIKANSKFIGITEL. MAPS structures are depicted in Figure 7 and Supporting Information.

The immunological activities of peptides can, however, be improved by constraining them to mimic protein substructures (*8*, *16*−*18*). This expedient though needs synthetic methods that can be used to broaden the approach (*17*). Recently, two of us reported a solid-phase synthetic method for replacing main-chain hydrogen bonds (NH → O=CRNH) with a hydrazone covalent mimic (NN=CH−CH_2CH_2) (*19*). On average, about 60% of the amino acids in globular proteins engage in main-chain hydrogen bonds (*20*). Since protein substructures can be distinguished by differences in hydrogen bonding patterns (*19*), we reasoned that the hydrazone link might be used to constrain peptides to different conformations depending on where it was positioned in the peptide (*19*). Indeed, the hydrazone link has been used to constrain short peptides as α-helices (*19*) and loops (*21*). It is also chemically stable in aqueous solution at physiological temperature and pH as required of immunogenic peptides (*19*).

Since the hydrazone link satisfied important criteria, we decided to test whether it could improve the antigenicity (affinity for antibodies) and immunogenicity (generation of neutralizing antibodies) of a short peptide from the V3 loop. The V3 loop is a 35-residue (± 2 residues) disulfide loop on gp120 that encompasses a neutralizing epitope on TCLA HIV-1 viruses (*22*). V3 peptides have provided important testing grounds for synthetic vaccine approaches (*23*−*27*). Interest in the V3 loop has also led to the identification of potent neutralizing V3 MAbs (*28*−*30*). We decided to use one of these MAbs to explore the antigenicity of hydrazone-linked peptides with the goal of identifying a constrained peptide mimetic of the MAb bound conformer using enhanced affinity as a measure of fit. The mimetic could then in turn be tested as an immunogen.

We relied mostly on MAb 58.2 (*30*) as a template for our studies. MAb 58.2 potently neutralizes TCLA HIV-1_{MN} and binds a V3 epitope that encompasses a short neutralizing determinant (SND), IGPGRAF, that is shared by the majority of Clade B HIV-1 viruses (*22*). X-ray crystallography of Fab 58.2-peptide complexes were being pursued and became available during the course of this project (*21*). Essentially, we found that constraining a short V3 peptide with the hydrazone link enhanced its affinity for MAb 58.2. This identified a cyclic peptide that showed considerable mimicry of the MAb 58.2 bound conformer by NMR spectroscopy. In turn, we found that the mimetic was a far better immunogen than the corresponding unconstrained peptide demonstrating the effectiveness of this approach.

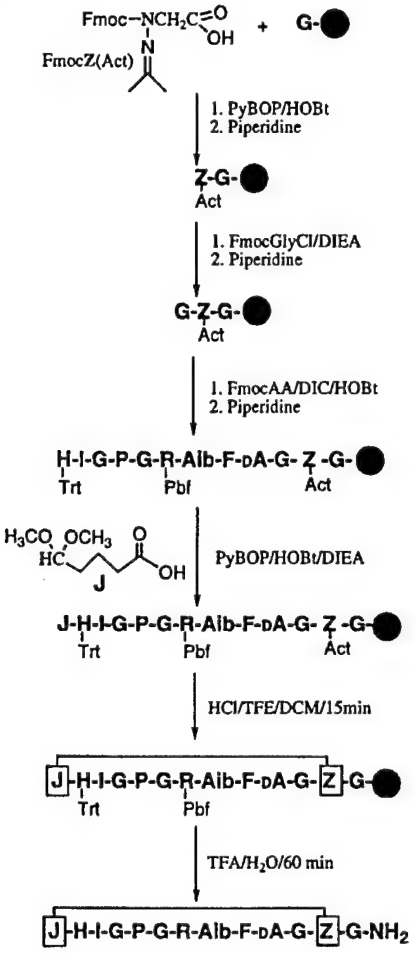
MATERIALS AND METHODS

Peptide Synthesis. Linear and cyclic peptides were synthesized by solid-phase methods. Side chains were protected using Fmoc-His(Trt), Fmoc-Arg(Pbf), Fmoc-Cys(Trt), Fmoc-Lys(Boc), and Fmoc-Lys(Fmoc). Peptides were purified with Gilson analytical and preparative HPLC systems using linear gradients formed from solvent systems A: 0.1% TFA/H_2O (v/v) and B: 0.1% TFA/AN (v/v).

Linear peptides (Table 1) were synthesized with an ACT-350 multiple peptide synthesizer (Advanced Chem Tech) on Rink amide resin (0.4−0.6 mmol/g) using standard Fmoc synthesis with DIC/HOBt coupling. The N-termini of peptides were acetylated on the resin with 15% acetic anhydride in DMF. The acetylated peptides were cleaved with Reagent K (*31*), precipitated with diethyl ether and purified using a preparative RP C-18 column (Cosmosil 5C18-AR, 2 × 25 cm) eluting at 8 mL/min with a 0−50% AN gradient over 30 min. Peptides were confirmed by FAB MS or ES MS.

Cyclic hydrazone-linked peptides (Table 1) were synthesized by solid-phase synthesis on Rink amide resin (0.4−0.6 mmol/g) according to Scheme 1 using previously described procedures (*19*). These syntheses required the prior preparation of J and Fmoc-Z(Act) (*19*). *N*-α-Fmoc-glycine chloride was prepared according to Carpino et al. (*32*) and shown to be >98% acyl chloride by the methanol test (*32*). Initially, Fmoc-GZ(Act)G/C(Trt)-resin was synthesized manually in resin packets (*19*). The most critical step was the coupling of Fmoc-Gly chloride to Z(Act). Fmoc-Gly chloride was freed of trace HCl under vacuum before use, mixed with

Scheme 1



NMP, and briefly shaken with Z(Act)G/C(Trt)-Resin before adding DIEA (*19*). The product, Fmoc-GZ(Act)G/C(Trt)-resin, was removed from the resin packet and transferred to the ACT-350 peptide synthesizer where it was extended using standard Fmoc synthesis. The addition of Fmoc-Arg(Pbf) to Aib was the only difficult coupling step, proceeding ≈60% despite triple coupling. After adding the final amino acid, Fmoc-His(Trt), the peptide-resin was transferred to resin packets, the Fmoc group removed with piperidine, the peptide capped with J and then rapidly cyclized in acidic 20% TFE/DCM (*19*). Cyclic peptides were cleaved from the resin and deprotected with 5% H_2O/95% TFA and precipitated with diethyl ether. Thiol scavengers were excluded from the cleavage reaction since high concentrations of free thiol in TFA destroy the hydrazone link (*19*). The cyclic peptides were typically purified by HPLC using a preparative RP C-18 column (Cosmosil, 5C18-AR) eluting at 8 mL/min with a 0−20% AN gradient over 5 min followed by 20−40% AN over 30 min (Supporting Information). The yields of cyclic hydrazone-linked peptides (based on resin loading) were ≈5−10%. Each purified cyclic peptide showed one peak by HPLC and was confirmed by FAB MS or ES MS.

The disulfide loop 6 (Table 1) was initially prepared as a linear peptide using Fmoc-Cys(Trt) and cleaved from the resin with Reagent K as described above. The linear peptide (1 mg/mL) was quantitatively cyclized in 20% DMSO/H_2O, pH 4.0, in 1 day (*33*). After confirming the loss of free thiol using Ellman's test (*34*), the reaction mixture was lyophilized and loop 6 was purified by HPLC on a RP-C18 column and confirmed by ES MS.

MAPS Synthesis. MAPS (*35*) were modified (*36*, *37*) and prepared for immunizations using a chemical ligation protocol (*37*). First, a chloroacetylated MAPS core (Table 1) was synthesized on TGR resin (NovaBiochem, 0.1 mmol/g) with the ACT-350 synthesizer using Fmoc synthesis. *N*-α-Fmoc-Lys(Fmoc) was used at the branch points. The N-termini of the four branches of the MAPS core were chloroacetylated on the resin with 10 equiv of chloroacetic acid anhydride in DMF for 15 min. The chloroacetylated MAPS core was cleaved from the resin and deprotected by treating it with 5% H_2O/95% TFA for 1 h. It was then precipitated with cold diethyl ether, lyophilized, and purified by HPLC on a preparative RP C4 column (Vydac, 214TP1022, 2.2 × 25 cm) eluting at 15 mL/min with a 0−20% AN gradient over 10 min followed by a 20−30% AN over 20 min. Pure chloroacetylated MAPS core (5% yield) was identified by ES MS: calcd (MH+), 9377, obs 9374 (Supporting Information).

Second, linear 5/C or loop 5/C (Table 1) were chemically ligated to the chloroacetylated MAPS core. A 5-fold molar excess (relative to chloroacetylated arms) of linear 5/C (2.8 mg) or loop 5/C (2.7 mg) was mixed with chloroacetylated MAPS core (1.0 mg) in 0.5 mL degassed 0.5 M Tris, pH 8.9, and 3 M guanidinium·HCl containing 8 mM EDTA under nitrogen. The reactions were monitored on an analytical RP C4 column (Vydac, 214TP54) eluting at 2 mL/min with a 0−20% AN gradient over 10 min followed by a 20−32% AN over 30 min (Supporting Information). The reaction appeared complete after 2 h but was continued overnight to ensure a complete reaction. The products were purified on the preparative RP C4 column (Vydac, 214TP1022) eluting at 15 mL/min with a 0−20% AN gradient for 5 min followed by a 20−35% AN for 30 min. These purifications yielded 0.8 mg of linear 5-MAPS (52% yield based on MAPS core) and 0.6 mg of loop 5-MAPS (40% yield) which were stored dry at − 20 °C. Both MAPS were soluble in water. Linear 5-MAPS was identified by ES MS: calcd (MH+), 14244, obs 14245. Loop 5-MAPS was identified by MALDI MS: calcd (MH+), 14228, obs 14223 (Supporting Information).

Immunization. Linear 5-MAPS and loop 5-MAPS were prepared in separate lots for immunizations by dissolving them in deionized water and vigorously mixing with an equal volume of twice the recommended concentration of RIBI adjuvant system R-730 (RIBI ImmunoChem Research, Inc) in PBS. Each rabbit (NZW female, 2−2.5 kg) was immunized with 1 mL of emulsion at multiple sites (subcutaneous, intraperitoneal, intradermal, intramuscular) as recommended by RIBI on days 0 (200 μg MAPS), 28 (100 μg MAPS), 71 (50 μg MAPS), and 225 (50 μg of linear 5-MAPS or 17 μg for loop 5-MAPS). The rabbits were bled on days 0, 28, 38, 44, 71, 82 (week 12), 89, 225, and 235 (week 34).

ELISA Titers. Titrations of rabbit antisera against linear 5, loop 5, and rgp120 were carried out in duplicate, and the results were averaged. Background absorption was determined by titrating each antiserum in the same manner but without antigen.

Linear 5/C and loop 5/C (Table 1) were conjugated to biotin-BMCC (Pierce) according to the manufacturer. Bio-

tinylated-peptide (10 ng/100 μL of PBS/well) was added to 96-well neutrAvidin coated microtiter plates (Pierce) for 2 h at room temperature which were then washed four times with distilled water. Rabbit antisera was serially diluted 2-fold with 1% BSA/0.05% Tween 20/PBS into titer wells and incubated for 1 h at 37 °C. After washing 10 times with distilled water, bound antibody was determined using alkaline phosphatase-conjugated goat antirabbit IgG F(ab′)$_2$ (Pierce) [100 μL/well of 1/500 dilution of IgG F(ab′)$_2$ in 1% BSA/0.05% Tween 20/PBS]. After incubation at 37 °C for 1 h and washing 10 times with distilled water, 100 μL of developing solution (1 mg of pNPP/mL of developing buffer) was added. The developing solution was prepared by adding 5 mg of pNPP (Sigma) to 5 mL of developing buffer (50 mL of diethanolamine, 50 mg of $MgCl_2 \cdot 6H_2O$, and 97.5 mg of NaN_3 were brought to 500 mL, pH 9.7, with ddi·H_2O and NaOH). Optical density was determined at 415 nm after 30 min with a Molecular Devices SpectraMAX 190.

Titers to protein were determined using rgp120$_{MN}$ (ImmunoDiagnostics Inc.). High binding 96-well microtiter plates (Costar, polystyrene 1/2 area) were coated overnight at 4 °C with sheep PAb specific for the C-terminus of rgp120 (no. 6205, International Enzymes, Fallbrook, CA, 0.5 μg of PAb/50 μL of PBS/well). After removal of unbound antibody by washing four times with distilled water, rgp120 (0.05 μg of rgp120/50 μL of 1% BSA/0.05% Tween 20/PBS/well) was added and incubated at room temperature for 2 h. Unbound rgp120 was removed by washing four times with distilled water, antibody was added, and the assay was continued in the same manner as described above.

End-point titers are the highest dilution of antisera that gave >2× background absorbance. Midpoint titers of antibodies for rgp120 are equal to the dilution of antisera that gave an optical density of 1.0 after correcting for background.

Competition ELISA Assay. Each experiment was carried out in duplicate and the results were averaged. First, linear 5/C-biotin (10 ng/well) or rgp120 (0.025 μg of rgp120/well) were bound to titer wells as described above. A subsaturating concentration of antibody was premixed with 4-fold serially diluted competing peptide and transferred to the antigen-coated wells (50 μL/well in 1% BSA/0.05% Tween 20/PBS). After incubation at 37 °C for 1 h, the unbound antibody was removed by washing 10 times with distilled water. Quantification of the bound antibody was carried out with alkaline phosphatase-conjugated goat antirabbit IgG as described above. 100% binding to plate-adsorbed antigen was determined in the absence of competing peptide. Background absorbance was determined without plate-adsorbed antigen and subtracted from the average of values with rgp120.

Neutralization Assays. Neutralization assays were carried out in triplicate in 96-well tissue culture plates (Costar). In a typical assay, 100 $TCID_{50}$ virus (MN or JR-CSF) was mixed with 2-fold serially diluted rabbit antisera or MAb in 100 μL of culture medium and incubated for 1 h at 37 °C. Then either 2×10^4 human T-lymphoid H9 cells (to MN) or 5×10^4 phytohaemagglutinin-stimulated, peripheral blood mononuclear cells (to JR-CSF) were added to each well and incubated for 24 h at 37 °C; PBMCs were pooled from three donors. After washing three times with culture medium and bringing each well to 200 μL with culture medium, incubation was continued for 7 days at 37 °C under 5% CO_2. Then 90 μL from each well was mixed with 10 μL of 10% empigein (Calbiochem, La Jolla, CA) and assayed for p24 in an ELISA (*38*). Reported values from the ELISA are based on median values from the triplicate determinations for each concentration of antibody.

NMR Spectroscopy. Peptides were dissolved in 0.5 mL of 10% D_2O/90% H_2O (v/v) and adjusted to pH 5.0 with aqueous NaOH and HCl solutions. Dioxane was added as an internal chemical shift reference (3.75 ppm). 1D and 2D COSY, P. E. COSY, TOCSY, and ROESY experiments were carried out with a Bruker AMX 500 spectrometer. 1D NMR spectra were run with presaturation of water. 2D ROESY spectra were run in the phase-sensitive mode using TPPI for quadrature detection in f_1. Typically, 32 transients of 1024 data points were acquired in the f_2 dimension at a spectral width of 5555 Hz with 512 t_1 increments which were zero filled to 2024 data points. A 400-ms mixing time was used. Additional ROESY spectra were acquired for loops 5 and 7 at 10 °C using a Bruker DRX-600 MHz spectrometer at a spectral width of 7186 Hz using eight transients for each of 512 increments and a 400-ms mixing time. Temperature coefficients were obtained from a series of 1D spectra. Probe temperatures were determined using a methanol standard. 1D and 2D data were processed and analyzed on a Silicon Graphics Indigo2 Extreme computer using Felix (Molecular Simulations, Inc).

RESULTS

MAb 58.2. MAb 58.2 was generated by White-Scharf et al. (*30*) to a 40-residue MN peptide (RP70) that encompassed the V3-disulfide loop. It was the most potent neutralizing MN MAb among >3000 V3 positive hybridomas (from 85 mice) generated by the peptide. MAb 58.2 also weakly neutralized primary isolates (*30*). MAb 58.2 is about twice as potent against MN as MAb 50.1 from the same series (*30*) that has been grouped with the most potent MN neutralizing human V3 MAbs (*39*). Although MAb 58.2 is technically an antipeptide MAb, its high potency suggested that it bound the MN V3 epitope in a conformation that could be used to represent the native structure.

Several approaches have been taken to characterize the MAb 58.2 epitope (*21*, *30*, *40*−*42*). Phage-displayed peptides and peptide arrays identified GPxR, the least variable region of the SND, as the core epitope (*40*). This specificity most likely accounts for the broad recognition of gp120s from primary isolates by MAb 58.2 (*41*). Alanine scanning of a RIHIGPGRAFY peptide with MAb 58.2 showed that it binds IGPGRAF (*30*). Later, serial deletion competition ELISAs indicated that MAb 58.2 binds RIHIGPGRAFY with the terminal amino acids contributing about 1000-fold to affinity (*42*).

A crystal structure of a Fab 58.2−Aib peptide complex showed that the antibody binds RIHIGPGR(Aib)FY (*21*). The bound structure (Figure 1) has been characterized as a strand followed by three overlapping β-turns (*21*). It can also be viewed in terms of a strand−β turn−pivot−β turn motif. Parsing the conformer in this manner identified structural elements that were useful for the design, manipulation, and conformational analysis of a mimetic. In this rendition, Fab 58.2 binds RIHI as a strand, and GPGR and R(Aib)FY as overlapping type I [$\alpha_R \rightarrow \alpha_R$] β-turns. Both β-turns are

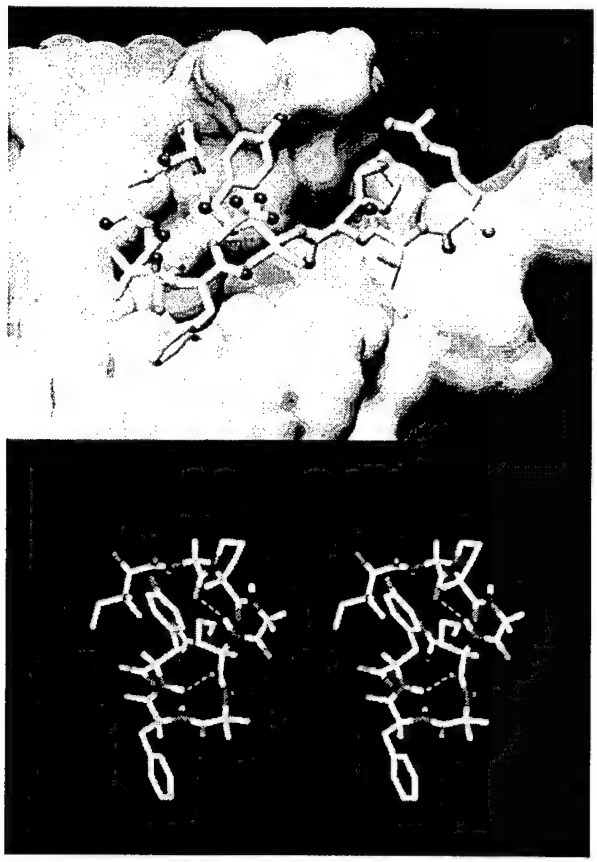


FIGURE 1: (top) Fab 58.2 bound MN V3 Aib-peptide, RHIG-PGRAibFY (*21*), PDB accession code 1f58. V3 Aib-peptide (stick model) against a Connolly surfaced Fab 58.2 (light chain, pink; heavy chain, blue) as rendered by Mike Pique using AVS. (bottom) Stereoview of Fab 58.2 bound, IGPGR(Aib)FY.

Table 2: Comparison of the Conformations Adapted by GPGRAFY Peptides and Analogs in Aqueous Solution and Fab-Peptide Crystal Structures in Terms of a β-Turn−Pivot−β-Turn Motif

structure MN sequence	β-turn GPGR	pivot[a] R	β-turn RAFY
linear Aib-peptide NMR structure[b]	random coil	β	random coil[c]
loop 5 NMR structure[b]	type I	α[d]	type I
Fab 58.2-V3 Aib-peptide (*21*)	type I	β	type I
Fab 59.1-V3 peptides (*44*, *45*)	type II	α	type I
Fab 50.1-V3 peptide (*43*)	type II[e]	NB[f]	NB[f]

[a] Arg is shared by both β-turns and acts as a pivot occupying either the α- or β-region of Ramanchandran space. [b] From this study. Dynamic equilibria of structures including the indicated structures. [c] May include a minor population of β-turn. [d] It is likely that R also occupies the β-region, but overlapping NOEs precluded this assignment. [e] Densities for GP and part of G6 (ϕ angle) were detected and agree with a type II turn. [f] Not bound by Fab.

characterized by main-chain i, i + 3 hydrogen bonds and share Arg that acts as a pivot to orient the turns with respect to one another in a double-headed loop (Figure 1). This motif also provides a common framework for comparisons with the conformations of V3 peptides bound by two other MN neutralizing antibodies, Fabs 50.1 (*43*) and 59.1 (*44*, *45*) (Table 2), that are referred to in this study.

Design and Synthesis. Since a Fab 58.2-peptide crystal structure was not initially available, we began with a hypothetical structure based on a secondary structure prediction made by LaRosa et al. (*22*) who predicted that the V3

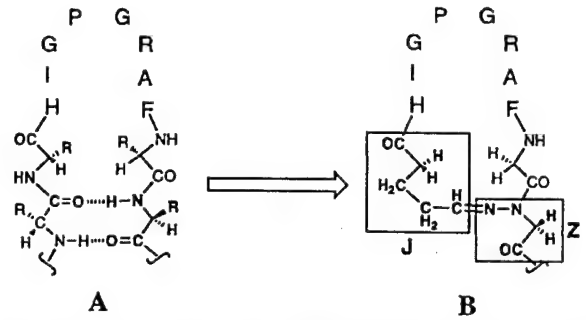


FIGURE 2: Transformation of a predicted MN V3 β-hairpin structure (A) into a hydrazone-linked loop 1 (B). The hydrazone link (boxed areas) was formed in a reaction between J and Z (Scheme 1).

sequence would form a β strand−type II β turn−β strand−α helix with GPGR occupying a type II turn. One manifestation of this prediction would place the SND in a β-hairpin. It was recently reported that a related V3 IIIB peptide binds a IIIB neutralizing MAb as a β-hairpin (*46*). β-hairpins are characterized by various sized loops at their apexes and a stem (*47*). The stems consist of antiparallel β-strands that hydrogen bond across alternating pairs of amino acids. The hydrogen bonds provide sites for substitution with the hydrazone link (*19*). For example, a predicted β-hairpin for the V3 MN sequence that includes the SND and a pair of hydrogen bonds characteristic of a stem is shown in Figure 2. The hydrazone link can be used to replace i + x → i main-chain hydrogen bonds in the manner shown for this structure to give in this case, loop 1 (Figure 2). However, the initial pair of hydrogen bonds could traverse a different set of amino acids determined by the size and position of the loop.

A set of corresponding linear and cyclic peptides was synthesized beginning with linear 1 and loop 1 (Table 1 and below). Loop 1 was then expanded using enhanced affinity for MAb 58.2 to identify a complementary loop while limiting as far as possible the composition of the loop to the SND. Corresponding linear and cyclic peptides were used to assess the effect of cyclization on affinity, independent of changes in amino acids. The hydrazone-linked peptides were synthesized on Rink amide resin using manual and machine-assisted multiple peptide synthesis according to Scheme 1. The coupling of Fmoc-amino acids to Z requires the use of an Fmoc-amino acid chloride since the reaction is sterically hindered (*19*). Fmoc-Gly chloride was used for coupling to Z for this series. Coupling to Z is not necessarily restricted to Gly (*19*). It required little more effort to assemble a hydrazone-linked peptide than a linear peptide and the yields were high for cyclic peptides of this complexity. Individual amino acids and their replacements are referred to using the numbering system in Table 1.

Peptide Affinities. The affinities of linear and cyclic peptides for MAb 58.2 were determined using a competition ELISA assay (Figure 3) and are reported as affinity indices, EC_{50} values of the peptides relative to the EC_{50} value for linear 1 (Table 3). Assays were carried out at 37 °C to simulate the conformational differences that occur at physiological temperature.

Loop 1 bound MAb 58.2 with an affinity index of 23. A Fab 58.2−loop 1 crystal structure has been reported (*21*). The propylene moeity of the hydrazone link appears fully

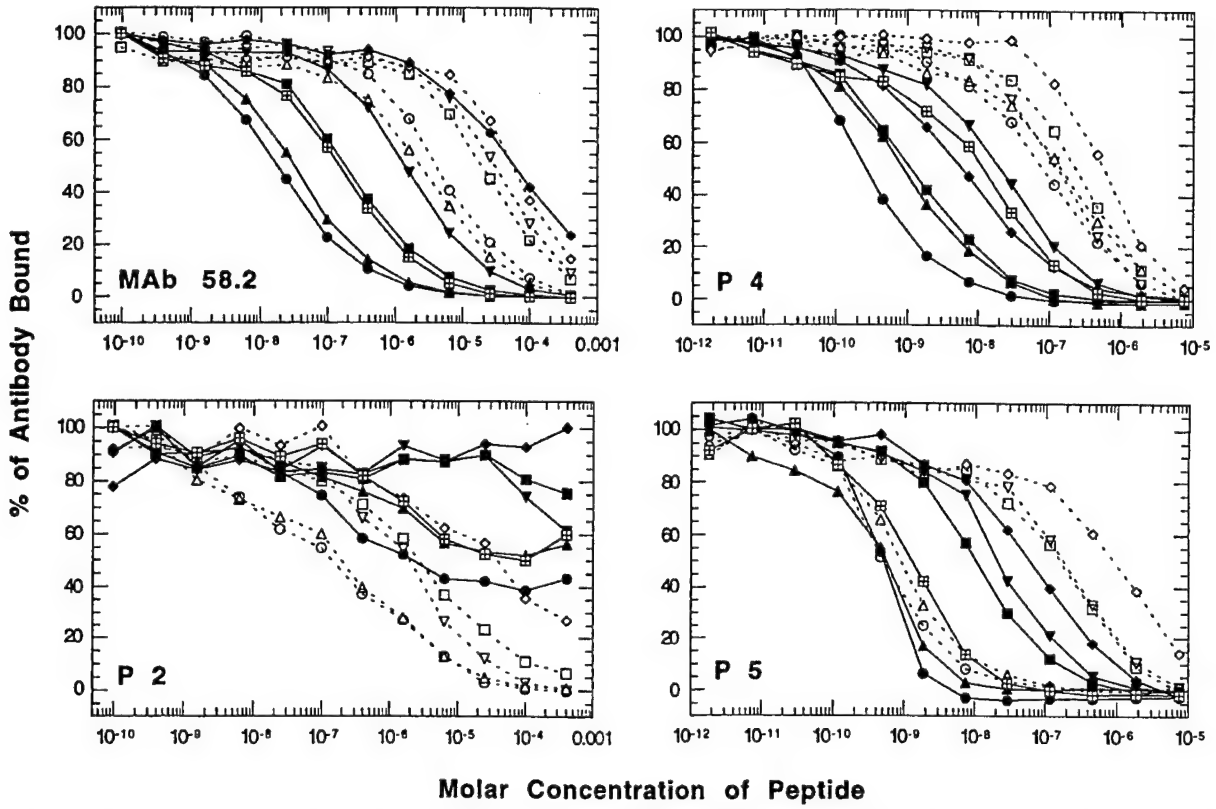


FIGURE 3: Competition ELISA assays of linear (1, ▽; 2, □; 3, ◇; 4, △; 5, ○) and loop peptides (1, ▼; 2, ■; 3, ♦; 4, ▲; 5, ●; 6, square with cross) for MAb 58.2 and rabbit antisera P2 (anti-linear 5), P4 and P5 (anti-loop 5). Peptides compete with plate-bound rgp120 (MAb 58.2, P4, P5) or linear 5 (P2). Percent antibody bound refers to the antibody bound by the plate-bound antigen. The peptides are identified in Table 3 which also includes relative EC_{50} values from these data.

Table 3: Affinity Indices of Linear and Loop Peptides for MAb 58.2 and P4 and P5 rgp120-Specific Antibodies[a]

peptide	sequence	MAb 58.2	P4	P5
linear 1	Ac-GHIGPGRAFGGG-NH_2	1.0	1.0	1.0
linear 2	Ac-GHIGPGRAFGGGG-NH_2	1.5	0.6	1.2
linear 3	Ac-GHIGP(D-Ala)RAFGGGG-NH_2	0.3	0.2	0.2
linear 4	Ac-GHIGPGR(Aib)FGGGG-NH_2	13	1.0	222
linear 5	Ac-GHIGPGR(Aib)F(D-Ala)GGG-NH_2	10	1.4	400
loop 1	[JHIGPGRAFGZ]G-NH_2	23	6.5	9.5
loop 2	[JHIGPGRAFGGZ]G-NH_2	167	108	20
loop 3	[JHIGP(D-Ala)RAFGGZ]G-NH_2	0.5	22	3.3
loop 4	[JHIGPGR(Aib)FGGZ]G-NH_2	968	144	370
loop 5	[JHIGPGR(Aib)F(D-Ala)GZ]G-NH_2	1667	464	400
loop 6	Ac-[CHIGPGR(Aib)FGGC]-NH_2	176	11	167

[a] Affinity indices are reciprocals of the (EC_{50} values for the peptide)/(EC_{50} value for linear 1). The EC_{50} values were obtained from Figure 3.

extended in the Fab 58.2 bound structure suggesting that loop 1 is barely wide enough to fit the binding site. This impression was strengthened by the later observation that Gly10 in loop 1 was rotated relative to Tyr10 in the Fab 58.2 bound linear Aib-peptide (Figure 1). When loop 1 was expanded by adding a second Gly to the C-terminal end of the SND to give loop 2, the affinity index improved to 167. We therefore eliminated smaller loops from consideration. When loop 1 was instead expanded from the N-terminus by adding Ileu to give [JIHIGPGRAFGZ]G-NH_2, the affinity index remained at about 23 (not shown). Since the higher affinity index for loop 2 suggested considerable complementarity with the MAb 58.2 binding site, ring expansion was stopped to limit the mimetic as far as possible to the SND. For comparison, MAb 58.2 binds RP70 (complete MN V3 disulfide loop peptide) with 22-fold higher affinity than an MN linear peptide that encompasses the complete MAb 58.2 epitope (*42*).

Three additional modifications were made to loop 2 at the **X** positions, [JHIGP**X**R**X**F**X**GZ]G-NH_2, to either probe conformational preference and/or improve the affinity of loop 2 for MAb 58.2. (i) Since Fab 58.2 binds GPGR as a type I turn (Figure 1), loop 3 was prepared with a Gly6/D-Ala substitution as a potential probe for this preference. D-Ala at the i + 2 position of a β-turn in cyclic pentapeptides destabilizes the type I turn and favors the type II turn (*48*). The affinity index for loop 3 decreased from 167 to 0.5 indicating the difficulties of binding GP(D-Ala)R as a type I turn. An examination of the Fab 58.2−peptide crystal structure (Figure 1), indicated that the D-Ala α-methyl group would not interfere sterically if loop 3 were bound as a type I turn. Consequently, loop 3 provides a sensitive probe for

type I turn specificity. (ii) Fab 58.2 also binds R(Aib)FY as a type I turn (Figure 1). Aib favors the α_R-region of Ramachandran space (*49*) which is adopted by Aib8 ($\phi = -50°$, $\psi = -42°$) in the Fab 58.2 bound peptide (*21*). The additional α-methyl group on Aib (as compared with Ala) projects into space and does not contact Fab 58.2 in the crystal complex (Figure 1). When Ala8 in loop 2 was replaced with Aib to give loop 4, the affinity index improved about 6-fold to 968 indicating a conformational stabilization by Aib. (iii) The conformational preference of MAb 58.2 for Gly10 was probed. Gly is a preferred amino acid at the i, i + 3 position of type I turns, often occupying the α_L-region of Ramachandran space (*59*). Consequently, it is not surprising to find that D-amino acids can occupy i, i + 3 positions in type I turns in cyclic pentapeptides (*48*, *50*). When Gly10 was substituted with D-Ala to give loop 5, the affinity index improved to 1667. However, when D-Ala10 in loop 5 was replaced with Aib, the replacement had little effect on affinity (not shown) suggesting that X10 can occupy different regions of Ramachandran space in the bound state.

A disulfide loop 6 (Table 1) was prepared to determine the effect of a different linker on affinity. Disulfide links are often found just beyond β-strand pairs where they are offset diagonally from the strands (*51*). The disulfide loop 6 links cysteine side chains to form a 38-member ring while the hydrazone-linked loop 5 is a main-chain linked 37-member ring. Replacing the hydrazone link with the disulfide link reduced the affinity index of loop 6 for MAb 58.2 to 176 (Figure 3) suggesting that the hydrazone link can better accommodate the conformer bound by MAb 58.2 than the disulfide link.

The affinities of linear and loop peptides for a second V3 MAb 59.1 were compared to determine whether MAb 59.1, which binds peptides in a related but different conformation to MAb 58.2 (Table 2), could discriminate among these peptides. MAb 59.1 bound loops 2 and 4 with about 4-fold higher affinity than the corresponding linear 2 and linear 4 peptides (not shown). The substitution of Aib for Ala8 had little or no effect on affinity. Consequently, loop 5 stabilizes or can adapt to the MAb 59.1 bound conformer but not nearly as well as to the conformer bound by MAb 58.2.

One further comparison was made of the affinity of MAb 58.2 for the MN peptide, acetyl-RIHIGPGRAFY-NH_2, with a corresponding peptide, acetyl-SIHIGPGRAFY-NH_2, representative of the primary isolates that were weakly neutralized by the antibody (*30*). A competition ELISA assay showed that MAb 58.2 bound the MN peptide with 14-fold higher affinity than the primary isolate peptide (Supporting Information). The N-terminal Arg forms a bifurcated, electrostatic bond between the guanidinium side chain and an aspartyl carboxylate group on the Fab (Figure 1) that likely accounts for the higher affinity of the MN peptide.

NMR Structural Studies. The high affinity indices of loops 2, 4, 5, and 6 for MAb 58.2 (Table 3) suggested that these cyclic peptides adopted conformation(s) in water that mimicked the MAb 58.2 bound conformer to a much better degree than the linear peptides. To assess this possibility, NMR structural studies were carried out on loop 5 and compared with those for a linear V3 Aib-peptide, acetyl-GHIGPGR(Aib)FGGG-NH_2, in 10% D_2O/H_2O, pH 5.0, at or near ambient temperatures.

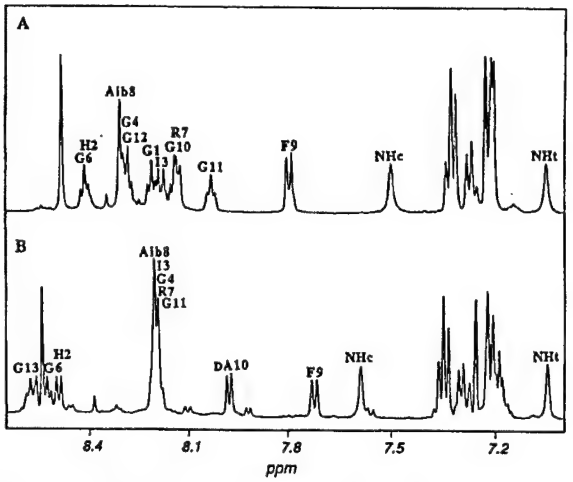


FIGURE 4: The amide proton region of 1D NMR spectra for (A) 20 mM linear Aib-peptide, acetyl-GHIGPGR(Aib)FGGG-NH_2, in 10% D_2O/90% H_2O at pH 5.0, 24 °C and (B) 20 mM loop 5 at pH 5.0, 21 °C, in 10% D_2O/90% H_2O.

A comparison of the amide NH regions of 1D NMR spectra for the linear Aib-peptide and loop 5 (Figure 4) revealed large shifts in resonance positions indicating major differences in chemical environment and conformation consistent with a structural basis for affinity enhancement. Both the linear and cyclic peptides showed a minor component (10%) that could arise from cis−trans isomerization at the Gly−Pro tertiary amide link. Another short, protected MN V3 peptide has been reported to isomerize in DMSO (*52*). As expected, the major component of loop 5 displayed a strong $G4^{\alpha}$-$P5^{\delta}$ NOE indicating a trans peptide bond (*53*). Fab 58.2 binds V3 peptides with Gly−Pro in the trans configuration (Figure 1) (*21*). Although we could not establish full connectivities for the 10% isomer, a $G4^{\alpha}$-$P5^{\alpha}$ NOE was evident for a minor component in loop 5 (not shown) indicating the presence of a cis Gly−Pro isomer (*53*). The acylhydrazone is another potential source of cis−trans isomerization. No G^{α}-Z^{α} NOE was apparent for the major component indicating that the main-chain C−N link is in the trans configuration relegating the -N−N- link to the "cis" position. Fab 58.2 binds loop 1 in this configuration (*21*). Loop 5 showed additional minor components (<5%) (Figures 4 and 6), which may reflect isomerization at the acylhydrazone or the presence of minor peptide impurities, possibly from racemization since no impurities were detected by mass spectroscopy (Supporting Information).

Further structural information was developed for the linear Aib peptide and loop 5 major components. The resonance positions for all protons for the linear Aib-peptide and loop 5 were assigned (Supporting Information) from 2D-NMR experiments using standard sequential assignment methods (*54*). These include stereospecific assignments for the loop 5 F9 β-methylene protons (Figure 5 and Supporting Information) which were made according to the method of Basus (*62*).

The overlap of several resonances for main-chain amide protons for loop 5 (Figure 4) prompted the synthesis of loop 7 (Ileu3 to Val3) which led to a small improvement in resonance separation (Supporting Information). Loop 7 showed a similar set of minor components. ROESY experiments for loops 5 and 7 displayed the same structure-defining

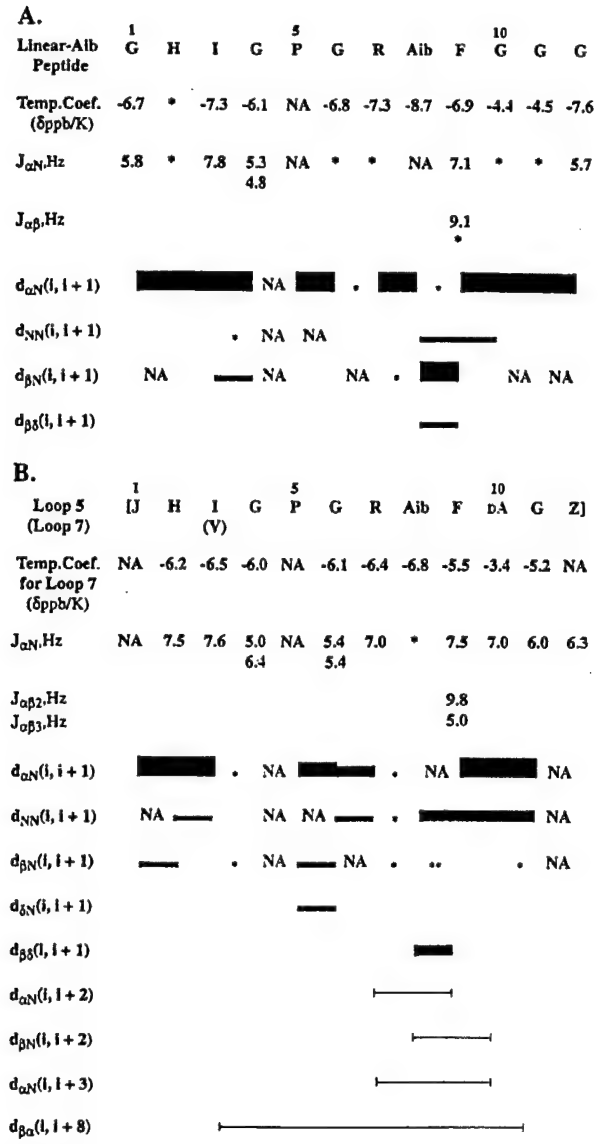


FIGURE 5: Summary of NMR data for the (A) linear Aib-peptide at pH 5.0, 24 °C and (B) loop 5 at pH 5.0, 21 °C. Temperature coefficients for loop 7 are included with the data for loop 5. Temperature coefficients were determined from linear regression analysis of five measurements made between 13−34 °C with $R >$ 0.996 (Supporting Information). $J_{\alpha N}$ and $J_{\alpha\beta}$ were determined from 1D and P. E. COSY spectra. Asterisk (*) indicates overlapping signals. Double asterisks (**) indicate that both Aib8$^{\beta}$-F9^{N} NOEs were medium intensity. Strong, medium-strong, medium, and weak NOEs are indicated by the heights of the filled bars. Lines are for weak NOEs except for the Aib8$^{\beta}$-(D-Ala)10^{N} NOE which was medium intensity. Blank spaces, no NOEs; NA, not applicable. The loop 5 F9$^{\beta 2}$ which resonates at 2.95 ppm showed a strong F9$^{N\beta 2}$ NOE (Figure 6) and a weak F9$^{\alpha\beta 2}$ NOE (not shown), while F9$^{\beta 3}$ at 3.22 ppm showed a weak F9$^{N\beta 3}$ NOE (Figure 6) and a medium F9$^{\alpha\beta 3}$ NOE (not shown).

NOEs (Supporting Information) indicating similar structures. The affinities of loop 5 and 7 for MAb 58.2 differed by no more than 2-fold consistent with similar structures (Supporting Information). Since loop 5 and loop 7 are chemically and structurally very similar, the additional NMR data for loop 7 resulting from better separation of resonance positions are used to supplement the data for loop 5.

Secondary structures display characteristic NOE patterns (*53*), which along with other data can be used to identify

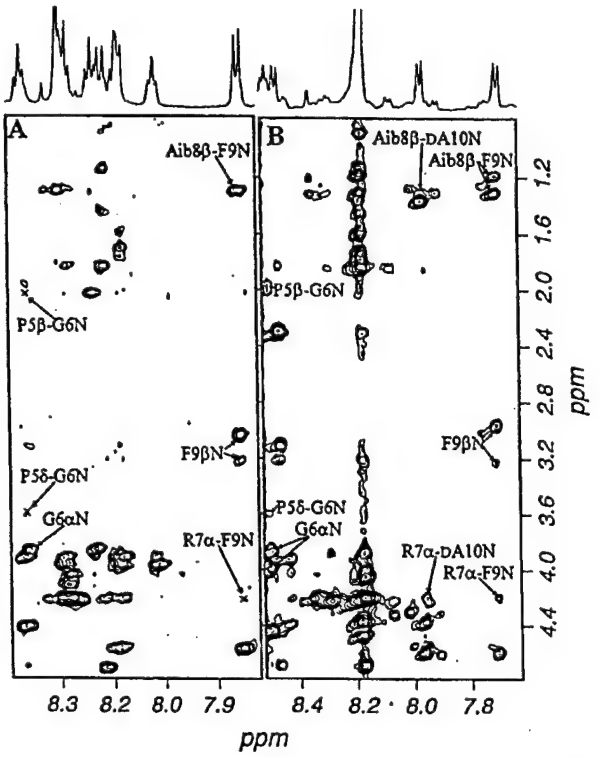


FIGURE 6: Portions of ROESY spectra for 20 mM solutions of (A) the linear V3 Aib-peptide, acetyl-GHIGPGR(Aib)FGGG-NH_2, in 10% D_2O/H_2O, pH 5.0, at 24 °C and (B) loop 5 in 10% D_2O/H_2O, pH 5.0, at 21 °C. NOEs that are indicative of β-turns are displayed by loop 5 but not by the linear Aib-peptide (x).

peptide conformations (*55*). The coupling constants, amide proton temperature coefficients, and NOEs for the linear Aib-peptide and loop 5 are summarized in Figure 5. Portions of spectra from 2D ROESY experiments identifying NOEs for the linear Aib-peptide and loop 5 are compared in Figure 6. Both peptides showed strong $d_{\alpha N}(i, i + 1)$ NOEs that are characteristic of unfolded forms (*55*). However, loop 5 displayed additional NOEs that indicated the stabilization of significant populations of β-turns within GPGR(Aib)F-(D-Ala).

Loop 5 showed P5$^{\delta}$-G6^{N}, P5$^{\beta}$-G6^{N} (Figure 6 and Supporting Information), and G6^{N}-R7^{N} (not shown) NOEs that were not observed for the linear Aib-peptide. In addition, the P5$^{\alpha}$-G6^{N} NOE for loop 5 was not as strong as for the linear peptide. These observations suggest that loop 5 stabilizes a type I β-turn at GPGR (*53*, *57*, *58*). The resonances for the G6 α-protons on loop 5, unlike those for the linear-Aib peptide, are separated (Figure 6) providing another indication of a better defined environment. Consequently, the $J_{\alpha N}$ coupling constants (5.4 Hz, 5.4 Hz) for the G6 α-protons may be significant. These coupling constants are, however, closer to those predicted for a type II turn (5 Hz) and a random coil (6.5 Hz) than a type I turn (9 Hz) (*58*). The medium-strong P5$^{\alpha}$-G6^{N} NOE could reflect the presence of a type II turn as well as a random coil (*53*, *57*, *58*). Type I and II turns are often but not always characterized by i, i + 3 hydrogen bonds (*56*, *57*). The temperature coefficient for the R7 amide proton was lower for loop 7 (6.4 ppb/K) than for the linear Aib-peptide (7.3 ppb/K) consistent with an i, i + 3 hydrogen bond. However, the effect is small and may reflect other phenomena. In this regard, R7 and Gln, which

are highly conserved by HIV-1 at GPG(R/Q) (*3*), are also favored at this position in type II turns for their ability to form side chain interactions with the i amino acid that substitutes for main-chain i, i + 3 hydrogen bonds (*59*). MAb 59.1 binds GPGR in this manner (*44*, *45*). Although the NMR data suggests that loop 5 stabilizes GPGR as a type I turn, it is likely to be in equilibrium with less ordered conformers and possibly other turn types.

The NMR data suggests that loop 5 stabilizes a second β-turn at R(Aib)F(D-Ala) as well. The data for R(Aib)F(D-Ala) which are characteristic of helices and β-turns (*53*, *58*) are consistent with the presence of type I and III β-turns. The type III turn [$\alpha_R \rightarrow \alpha_R \rightarrow \alpha_R \rightarrow \alpha_R$] has been merged with and classified as a type I turn [$\alpha_R \rightarrow \alpha_R$] since in both cases the β-turn defining i + 1 and i + 2 amino acids occupy the α-region of Ramachandran space (*60*, *61*). However, a type III turn is also one turn of a 3_{10}-helix that is characterized by additional NOEs (*53*, *58*) that were observed for loop 5.

The NMR data for R(Aib)F(D-Ala) can be summarized as follows. (i) Sequential $d_{NN}(i, i + 1)$ NOEs for (Aib)F(D-Ala)G (not shown) are characteristic of helices and the type III turn (*53*, *58*). (ii) The $R7^{\alpha}$-$F9^{N}$ NOE is weak and of similar intensity to the $R7^{\alpha}$-(D-Ala)10^{N} NOE (Figure 6 and Supporting Information) suggesting a 3_{10}-turn, i.e., type III turn (*53*, *58*). It is very unlikely that these weak NOEs or the weak $P5^{\delta}$-$G6^{N}$ described above arise from minor components buried under major component resonances since they are observed at two different temperatures for loop 5—at 10 °C (Supporting Information) and 21 °C (Figure 6)—indicating the same temperature dependencies as the major component resonances. In addition, these NOEs were observed for loop 7 (Supporting Information). (iii) An $Aib8^{\beta}$-(D-Ala)10^{N} NOE is predicted for type I and type III turns and observed as an overlapping NOE (Figure 6) for one of the α-methyl groups on Aib8. This NOE is clearly resolved for loop 7 (Supporting Information). The resonances for the two α-methyl groups on Aib8 are well-separated for loop 5 as compared with the linear Aib-peptide (Figure 6) consistent with a more structured environment. The pattern of $F9^{\alpha\beta}$ coupling constants, $F9^{N\beta}$ and $F9^{\alpha\beta}$ NOEs, which are summarized in Figure 5 and the accompanying legend, indicate that the conformation of the F9 side chain is largely restricted to the g+ conformation (χ angle = −60°) (*62*), in agreement with the Fab 58.2 bound Aib-peptide (Figure 1) (*21*). (iv) The temperature coefficient for the D-Ala10 amide proton of loop 7 (−3.4 ppb/K) is low, which is consistent with its participation in a hydrogen bond to the R7 carbonyl oxygen atom that is characteristic of type I and III turns. The temperature coefficient for the D-Ala10 amide proton in loop 5 which is well-resolved (Figure 4) is similarly low (not shown). The NMR data for R(Aib)F(D-Ala) suggests a very substantial population of type I turns that may be embedded in and/or in equilibria with helical-like turns. The linear Aib-peptide may stabilize minor populations of similar structures as suggested by weak $d_{NN}(i, i + 1)$ NOEs (not shown) for AibFG and a lowered temperature coefficient for the G10 amide proton.

Overall, the NMR data for loop 5 is consistent with a dynamic equilibrium of structures that include regions of disorder and conformers that stabilize type I turns at GPGR and/or R(Aib)F(D-Ala) that are not as evident in the linear

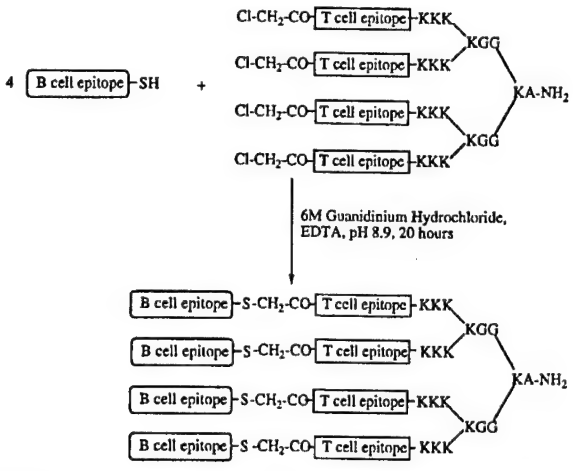


FIGURE 7: Synthesis of linear 5 and loop 5 MAPS. The B-cell epitope, either linear 5/C or loop 5/C, was reacted with a chloroacetylated T-cell epitope, QYIKANSKFIGITEL, on a MAPS core. The branch points were formed by coupling amino acids to the α- and ϵ-amino groups of Lys. Chemical structures are shown in Table 1 and Supplementary Information.

Aib-peptide. The enhanced affinity of loop 5 for MAb 58.2 which binds GPGR(Aib)FY as overlapping type I turns can thus be attributed to conformational stabilization. However, it remains unclear whether loop 5 stabilizes the precise conformation recognized by MAb 58.2 (Table 2) or whether further conformational adjustments are required for binding.

Immunogenicity. The corresponding linear 5 and loop 5 were incorporated into individual MAPS (Figure 7) for generating antibodies. The basic MAPS design (*35*) was modified by increasing the distance between branch points with tandem glycines (*36*) and inserting three tandem lysines at the carboxyl end of the four branches which imparts aqueous solubility (*37*). C-terminal cysteines on linear 5/C and loop 5/C were reacted with the chloroacetylated-tetanus toxoid T-cell epitope, QYIKANSKFIGITEL (TT2), on a four-branched MAPS core to give linear 5-MAPS and loop 5-MAPS (Figure 7). Chemical ligation opens a route to larger MAPS that can be synthesized with the purity required for characterization by MS (*63*). The TT2 epitope is capable of recruiting T-cell help in mice, rabbits, and humans (*64*, *65*). Both MAPS which are protein-sized (132 amino acids) were soluble in water and confirmed by MS (Supporting Information).

Two groups of three female NZW rabbits were immunized four times during the course of 34 weeks with linear 5-MAPS (rabbits P1−P3) and loop 5-MAPS (rabbits P4−P6) in RIBI's adjuvant (*66*). Dosage was minimized and the immunization schedule was lengthened to favor high affinity antibodies (*67*). High titers of antibodies to the immunizing peptide were observed following the third (12 week) and fourth (34 week) immunizations. Unless otherwise noted, analyses were carried out with the 12-week antisera.

The immunized rabbits showed striking differences in their antibody responses to the linear 5-MAPS and loop 5-MAPS (Figure 8). Each rabbit immunized with the loop 5-MAPS developed antibodies that bound rgp120, while rabbits immunized with the linear 5-MAPS did not (Figure 8 and Supporting Information). Both sets of rabbits developed high end-point titers (≥ 64 000 for the immunizing peptide) of antibodies that bound linear 5 and loop 5. However, while

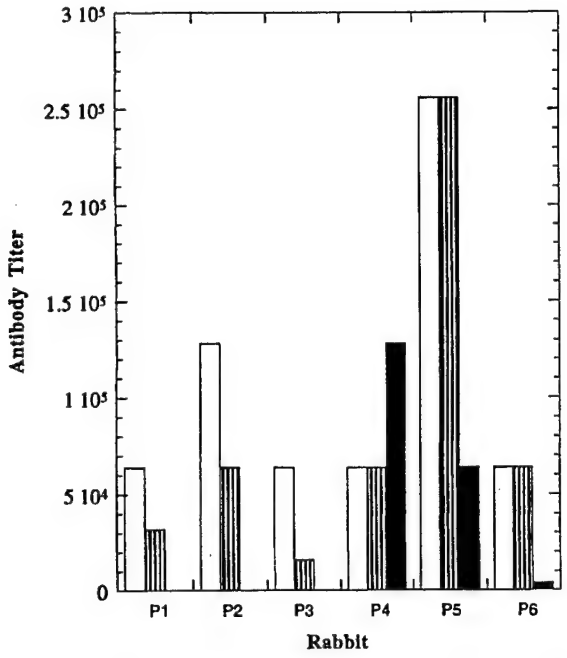


FIGURE 8: End-point titers of antibodies generated by rabbits P1−P3 (linear 5 MAPS) and rabbits P4−P6 (loop 5 MAPS) to linear 5 (clear), loop 5 (striped), and rpg120 (solid).

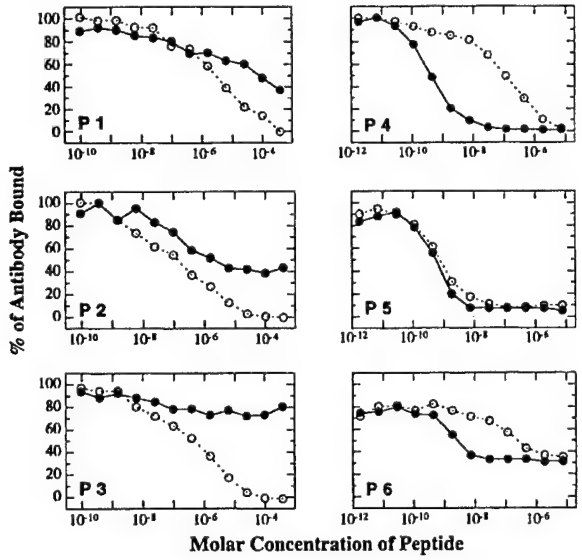


FIGURE 9: Competition ELISA assays of linear 5 (○) and loop 5 (●) for rabbit P1−P6 antibodies. Peptides compete with plate-bound linear 1 (P1−P3) or rgp120 (P4−P6). % Antibody bound refers to the antibody bound to the plate-bound antigen. The data for P2, P4, and P5 are from Figure 3.

the titer profiles for rabbits immunized with the linear 5-MAPS were similar, those immunized with the loop 5-MAPS showed distinct differences. P4 antisera showed higher titer to rgp120 than to linear 5 and loop 5. P5 antisera showed very high titers to linear 5 and loop 5 and relatively lower titer to rgp120. P6 antisera showed high titer to linear 5 and loop 5 and very low titer to rgp120.

The antisera were further evaluated in competition ELISA assays. The relative affinities of linear 5 and loop 5 for P1−P6 antibodies were compared (Figure 9). Since the P1−P3 antisera did not bind rgp120, a C-terminal biotinylated linear 5 peptide was adsorbed to neutrAvidin coated titer wells for competition ELISAs with P1−P3 antisera. The P1−P3 anti-linear 5 antibodies bound linear 5 with higher affinities than loop 5. A significant fraction of the P1 and P2 antibodies bound loop 5, while P3 antibodies showed a smaller population of antibodies binding loop 5. The competition curves for P1 and P2 are relatively flat reflecting heterogeneous populations of antibodies.

Competition ELISAs for the P4−P6 antibodies was carried out on plate-bound rgp120 to characterize the protein-reactive antibodies. The rgp120-binding P4−P6 antibodies showed very high affinities for loop 5 [$EC_{50} = (0.5-2) \times 10^{-9}$ M^{-1}, Figure 9]. The EC_{50} values are at least 10-fold lower than that shown by MAb 58.2 for loop 5 ($EC_{50} = 2 \times 10^{-8}$ M^{-1}, Figure 3). Also, the EC_{50} values for loop 5 for the P4−P6 antibodies are about 100−1000 times lower than the EC_{50} values for linear 5 for the P1−P3 antibodies [$EC_{50} = (0.2-2) \times 10^{-6}$ M^{-1}]. Although the P4−P6 rgp120-binding antibodies are heterogeneous, the individual antisera differed enough that distinctions can be drawn. The P4 antibodies showed the greatest similarity to MAb 58.2. P4 antibodies binds loop 5 with a 330-fold lower EC_{50} value than linear 5. In contrast, P5 antibodies bind loop 5 and linear 5 with very similar apparent affinities. The P6 rgp120-specific antibodies showed higher affinity for loop 5 but represent a minor population. The competition curves for P4−P6 are more nearly sigmoidal than those for P1−P3 indicating antibodies with similar affinities.

The preferences of P2 linear 5 antibodies and P4 and P5 rgp120-specific antibodies were compared with MAb 58.2 using full peptide scans (Figure 3). These scans provide a more complete picture of the differences in the responses of rabbits to linear 5 (P2) and to loop 5 (P4, P5) as well as animal-to-animal differences in the response to loop 5 (P4 vs P5). Relative EC_{50} values (affinity indices) for MAb 58.2, P4, and P5 rgp120-binding antibodies are compared in Table 3.

The P4 rgp120-specific antibodies came closest to duplicating the competition profile for MAb 58.2 (Figure 3). The affinity indices were similar with those for loop 5 > loop 4 > loop 2 > loop 1 > linear 4/5 > linear 1/2 > linear 3 (Table 3) for both P4 antibodies and MAb 58.2. The P4 antibodies also bound the disulfide loop 6 with higher affinity than the corresponding linear 4 peptide implying a conformational effect independent of the linker. Loop 3 is the only peptide that strongly deviates from the general agreement. Whereas the substitution of D-Ala for Gly6 in loop 2 to give loop 3 decreased the affinity index 334-fold for MAb 58.2, it reduced the index about 5-fold for the P4 antibodies (Table 3). MAb 58.2 appears unique among the antibodies examined in its ability to strongly discriminate against loop 3.

The P5 rgp120-specific antibodies showed strong preferences for Aib-peptides that appeared to be nearly independent of whether the peptide was cyclized or not (Figure 3). The affinity index of P5 antibodies for linear 4 (Aib-peptide) was 185-fold higher than that for the corresponding linear 2 (Ala-peptide) (Table 3) implying that Aib was a core amino acid and that the additional α-methyl group on Aib was specifically bound by the P5 antibodies. The similar affinities and sharper competition curves of Aib-peptides for P5 antibodies may reflect very high affinity, stoichiometric binding, of Aib-peptides to these P5 antibodies. Although P5 antibodies

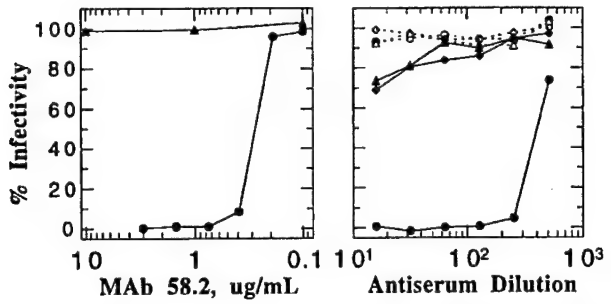


FIGURE 10: The neutralization of infection of H9 cells by 100 $TCID_{50}$ MN with (A) MAb 58.2 with (▲) and without (●) 41 μM loop 5 and (B) P1−P6 antisera, week 12 (P1, ○; P2, ◊; P3, △; P4, ●; P5, ♦; P6, ▲). The neutralization of MN by P4 antiserum is blocked by 41 μM loop 5 (not shown).

showed similar apparent affinities for linear and cyclic Aib-peptides, they showed a 17-fold higher affinity index for loop 2 (Ala-peptide) as compared with linear 2 (Ala-peptide). The higher affinity of gp120-reactive antibodies for a cyclic peptide was maintained, although it was not as great as observed for the P4 rgp120-specific antibodies (Table 3).

The P2 antibodies bound with higher affinities to linear peptides than to the corresponding cyclic peptides (Figure 3). The P2 antibodies showed about a 10-fold higher affinity index for linear Aib-peptides than for linear non-Aib peptides. Similar differences were observed for P1 and P3 antibodies (not shown). The apparent affinity enhancements of P1−P3 antibodies for linear Aib peptides (≈7−10-fold) were similar to MAb 58.2 (≈10-fold) suggesting a conformational effect. Regardless, Aib was not a core amino acid for most P1−P3 antibodies as it was for P5 antibodies (185-fold Aib effect). About half of the P2 antibodies bound loop peptides (Figure 3), while the remaining half did not. The inability to bind a loop could reflect conformational incompatabilities between anti-linear 5 antibodies and loops or it could reflect the requirement for a free N-terminal amino acid (wrap-around binding). However, at least a half of P1 and P2 antibodies and a quarter of P3 antibodies bound loop 5 in ELISAs (Figure 8) indicating that these antibodies do not require a free end and should not be limited by this requirement from binding rgp120.

The P4−P6 rgp120-binding antibodies share characteristics that were not found for the P1−P3 antibodies, namely, high affinities and a preference for cyclic peptides. Although the modified regions of the peptides could contribute to the affinity of rabbit antibodies for peptides (P5 antibodies for Aib), rgp120 is unmodified. The simplest interpretation of rgp120 binding is that the loop 5 antibodies are more compatible with the conformation of the V3 epitope on rgp120 than the linear 5 antibodies.

HIV Neutralization. Each of the rabbit antisera and MAb 58.2 were tested for the neutralization of MN infection of human T-lymphoid H9 cells (Figure 10). MAb 58.2 neutralized MN with IC_{90} = 0.75 μg/mL. P1−P3 linear 5 antisera had no effect on MN. The loop 5 P4 antiserum (12 week, 34 week) neutralized MN with titers of 128 and 100, respectively. P5 and P6 antisera (12 week) showed weak neutralizing activity at 1/16 antiserum dilution. Several controls were run concurrently. Neither preimmunization P4 serum, loop 5 alone (41 μM), MAb 58.2 mixed with loop 5 (41 μM), nor P4 antiserum (34 week) mixed with loop 5 (41 μM) neutralized MN.

Table 4: Antibody Midpoint Titers for rgp120, MN Neutralization Titers, and Potencies of MAb 58.2 and P1−P6 Antisera[a]

antibodies	rgp120 midpoint titer	neutralization titer IC_{90}	potency (rgp120 titer/ neutralization titer)
MAb 58.2	5.0×10^5	1.5×10^4	33
P1−P3 (12 week)	<125	<16	NA
P4 (12 week)	1.0×10^4	128	78
P4 (34 week)	1.0×10^4	100	100
P5 (12 week)	4.5×10^3	<16	>281
P6 (12 week)	200	<16	>13

[a] Antibodies are from 10-day bleeds following the third (12 week) and fourth (34 week) immunizations. rgp120 midpoint titers were determined from the reciprocols of MAb 58.2 (11 mg/mL) or antisera dilutions that gave an optical density of 1.0 in the standard ELISA (Materials and Methods). Neutralization titers are for reciprocols of the highest dilution of antibody that suppressed detectable levels of p24. Potencies are rgp120 midpoint titer/IC_{90}, neutralization titer. NA, not applicable.

Since a goal of this study was to determine whether the potent neutralizing activity of a template MAb could be recapitulated with a mimetic identified by the MAb, we compared the neutralizing potency of the rabbit antibodies with MAb 58.2 (Table 4). Potency is a measure of the concentration of an antibody required for neutralization. Since PAbs are complex mixtures of antibodies, we used midpoint antibody titer for rgp120/neutralization titer as a measure of potency (Table 4). These measures served two purposes. First, they provided a means for comparing the neutralizing activity of PAbs with a template MAb. Second, since the neutralization sensitivity of HIV-1 can vary (*68*, *69*), a MAb can provide a standard for wider comparisons.

MAb 58.2 and P4 antisera (12 week, 34 week) neutralized MN with similar potency (Table 4) indicating that loop 5 was capable of recapitulating the activity of MAb 58.2. Since MAb 58.2 compares favorably with the most potent V3 MAbs (see above), the P4 antiserum is also potent. On the other hand, the P5 rgp120-specific antibodies were less potent than MAb 58.2 (Table 4). The P6 rgp120-specific antibodies would have to be quite potent to account for even a weak neutralizing titer since they were generated at a far lower titer than the P4 and P5 rgp120-specific antibodies (Table 4).

Both MAb 58.2 and P4 antiserum were tested for the neutralization of JR-CSF, an HIV-1 primary isolate (*70*) that has proven refractory to other V3 MAbs (*71*). Although the JR-CSF V3 sequence, SIHIGPGRAFY, differs from the MAb 58.2 MN epitope, RIHIGPGRAFY, at the N-terminus of the MAb 58.2 epitope, MAb 58.2 weakly neutralized JR-CSF (IC_{50} = 200 μg/mL). This is in agreement with previous reports that MAb 58.2 weakly neutralized primary isolates (*30*). The neutralization of JR-CSF by MAb 58.2 was blocked by loop 5 (41 μM) indicating that neutralization was antibody-specific. Loop 5 (41 μM) by itself did not block infection. P4 antiserum (week 34) had no effect on JR-CSF infection at 1/4 dilution.

Clonal Selection. P4−P6 antibodies show very distinct characteristics indicating animal-to-animal differences in antibody responses to loop 5. First, the relative titers of P4−P6 antibodies to linear 5, loop 5, and rgp120 are distinctive (Figure 8). Second, the P4 and P5 rgp120-specific antibodies show clear differences in relative affinities for the same set

of linear and loop peptides in competition ELISAs (Figure 3). Third, P4 antiserum potently neutralized MN, while the P5 and P6 antisera were weakly neutralizing (Figure 10). These results imply a restricted clonal response of B-cells to loop 5 by the individual rabbits. The clearest case for clonal selection can be made with P5 rgp120-specific antibodies since they show a strong preference for the Aib-peptides. The Aib-peptides differ from the Ala-peptides by the addition of a single α-methyl group that represents about 1% of the mass of loop 5. However, most if not all of the P5 rgp120-specific antibodies show large affinity enhancements for the linear 4 Aib-peptide as compared with the corresponding linear 2 Ala-peptide (Figure 3). This implies that most of these P5 antibodies bind the extra methyl group suggesting that a very narrow population of rgp120-specific antibodies was generated by P5.

DISCUSSION

A striking result from the current study was the 5500-fold range of affinities of linear and loop peptides for MAb 58.2 (Figure 3). These affinity differences are most likely due to structural changes within the peptide rather than contributing or interfering interactions of modified regions of the peptide with MAb 58.2. The modifications were made exterior to the observed binding face of V3 peptides in Fab 58.2-peptide crystal complexes (*21*). The wide range of peptide affinities shows that the affinity of an antibody for a peptide can be extraordinarily sensitive to peptide structure.

Loop 5 binds >1000-fold more tightly to MAb 58.2 than the corresponding unconstrained V3 peptides (linear 1 and 2). Our literature search went back 27 years to the 17-residue disulfide loop of lysozyme to find a similarly large effect (*72*). The earlier study, as does the present one, proposed that the enhanced antigenicity and immunogenicity of the cyclic peptide was due to a conformational effect. This proposal was questioned for the lysozyme loop study (*73*) since no direct evidence was presented that the synthetic lysozyme loop adapted a native structure. The current study provides a structural analysis that supports a conformational basis for enhanced activity.

The relationship between the structure and affinity of a peptide for an antibody has been described by Anfinsen and co-workers (*12*). Formally, affinity can be viewed as the product of the fraction of the total peptide that is folded, reflecting K_{conf}, the equilibrium constant for folding, and the binding constant for the folded peptide. Anfinsen (*12*) calculated that K_{conf} for folding a 51-residue fragment of staphylococcal nuclease is 2×10^{-4} reflecting a 5000-fold affinity enhancement for the catalytically active form of the nuclease as compared with that of the inactive fragment for PAbs generated by the native protein (*12*). The >1000-fold improvement in affinity of loop 5 relative to the unconstrained V3 peptide for MAb 58.2 indicates that large affinity enhancements can extend to short constrained peptides (8-residues) as well. Whereas K_{conf} for the large nuclease fragment could reflect that for the reconstitution of a discontinuous epitope (topological assembly of amino acids distant in sequence), the latter result with loop 5 indicates that large effects can extend to continuous epitopes.

Since peptides are mostly disordered in aqueous solution (*11*), constraining peptides to mimic protein substructures would be expected to improve their affinities for native protein antibodies in many cases. Indeed, preliminary studies show that hydrazone-linked peptides designed to stabilize loops and α-helices detect antibodies in various antisera to different pathogens that go undetected with the corresponding unconstrained peptides (*74*, *75*). This has important implications since it suggests that the native protein antibodies that are most suited for serving as templates for synthetic vaccine development are being missed in screens with unconstrained peptides but can nonetheless be detected with constrained peptides because of their much higher affinities.

Loop 5 was also far more immunogenic than linear 5. This difference reflects structural differences in the peptides. Thus, the antibodies generated to linear 5 by the P1−P3 rabbits showed affinity preferences for linear peptides in comparison with loop peptides (Figures 3 and 9) indicating the difficulties of accommodating a loop structure. Neither could the P1−P3 antibodies bind rgp120 (Figure 8), despite exposure of the SND, implying that conformational incompatibilities prevented binding. On the other hand, the P4−P6 rabbits generated antibodies to loop 5 that bound rgp120 (Figure 8 and Supporting Information). These antibodies also bound loop peptides as well as or with higher affinity than linear peptides (Figures 3 and 9) demonstrating their capacity for binding a loop that also can account for why they bound rgp120.

The P4−P6 antibodies also neutralized MN (Figure 10), although with very different activities. Animal-to-animal differences in neutralizing activity generated by V3 peptides have been frequently observed (*23*−*25*, *27*, *76*). However, the bases for these differences have rarely been explored. In one study, a comparison of the amino acid specificities of antibodies in nonneutralizing and neutralizing IIIB antisera showed only minor differences (*24*). Consequently, the affinity profiles of P1−P6 antibodies for the linear and loop peptides (Figures 3 and 9) are striking for the differences they revealed.

Neutralizing titers are determined by the concentrations of neutralizing antibodies and their potencies. The concentrations of V3 antibodies were measured as titers for linear 5, loop 5, and rgp120 (Figure 8). Although titers were high to the immunizing peptide, the P4−P6 antisera showed quite different titers to rgp120. Added to this, the neutralizing potencies of the P4−P6 rgp120-specific antibodies also differed (Table 4). These measures together with the affinity profiles of the rgp120-specific antibodies (Figures 3 and 9) provide rationales that can account for differences in neutralizing activity.

The higher neutralizing titer of the P4 antisera was due to both the higher relative concentrations of P4 rgp120-specific antibodies (Figure 8) and their higher potencies (Table 4). The P4 rgp120-specific antibodies were similar in potency to MAb 58.2 (Table 4) demonstrating that the activity of a template MAb can be recapitulated with a PAb response. The P4 antibodies were also generated to a smaller MN epitope, HIGPGRAF, than MAb 58.2, which binds RIHIGPGRAFY. MAb 58.2 could derive a considerable affinity advantage (up to 1000-fold) from flanking amino acids (*42*) that cannot contribute to the affinity of P4 antibodies for MN. Thus the P4 response shows that potent neutralizing activity can also be focused onto a smaller epitope.

A comparison of the affinity profile of linear and cyclic peptides for the P4 rgp120-specific antibodies with the profile for MAb 58.2 showed strong similarities (Figure 3). This is consistent with similar fine structure and conformational specificities for the SND that could account for their similar potencies (Table 4). These profiles, however, also differ in some ways which as discussed below may reflect on requirements for neutralization via the SND.

The lower neutralizing titer of the P5 antisera relative to the P4 antisera was due more to the lower potency of the P5 rgp120-specific antibodies than to concentration differences (Table 4). It has been shown that the lowered potency of neutralizing MAbs for MN is synomynous with lower affinity for the virus (*79*). Differences in the affinity of antibodies for MN could arise from a number of sources, including differences in fine structure and conformational specificities as well as access to the epitope of the virus (*79*). The affinity profile of P5 rgp120-specific antibodies (Figure 3) shows a strong affinity enhancement for Aib-peptides (Table 3) compared with Ala-peptides, indicating that the added α-methyl group on Aib can contribute significantly to affinity. Consequently, the affinity of these P5 antibodies for MN could be suboptimal. The specificity of the P5 antibodies for Aib-peptides also indicates differences in the epitope recognized by P4 and P5 antibodies that could possibly lead to differences in accessibility to the MN epitope. Regardless, the observation that most if not all of the P5 antibodies show a specificity for Aib provides strong evidence for a very narrow population of P5 rgp120-specific antibodies.

The P6 antisera showed, as did the P5 antisera, lower neutralizing titer for MN (Figure 10). However, the bases for lower neutralizing titers of the P5 and P6 antisera differ. The P6 rgp120-specific antibodies were generated at a much lower titer than either the P4 or P5 rgp120-specific antibodies (Figure 8). Thus the lower neutralizing titer for P6 antisera reflects lower rgp120-specific antibody concentration rather than potency. The P6 rgp120-specific antibodies would have to be quite potent to account for even weak neutralizing titer (Table 4). The P6 rgp120-specific antibodies showed like the P4 antibodies a strong preference for loop 5 as compared with linear 5 (Figure 9) consistent with a similar conformational preference providing a basis for higher potency (Table 4). However, for the P6 rabbit, a third class of loop 5 antibodies predominated that bound both linear 5 and loop 5 but did not bind rgp120 (Figure 8).

The appearance of distinct populations of antibodies in the P4−P6 rabbits implies clonal selection and an amplification process. Amplification provides a mechanism for generating populations of antibodies with specificities for selected conformers or families of conformers. Jemmerson and Blankenfeld (*10*) showed that the activation of a B-cell for antibody generation is dependent on the immunogen achieving an affinity for the antigen receptor on the B-cell that exceeds that for binding the receptor, i.e., binding alone does not necessarily lead to antibody generation. The rgp120 specific antibodies generated by the P4−P6 rabbits showed much higher affinity for the corresponding immunogen, loop 5 ($EC_{50} \approx 10^{-9}$ M), than did the P1−P3 antibodies for linear 5 ($EC_{50} \approx 10^{-6}-10^{-7}$ M) (Figure 9), which may also be related to this phenomenon and the development of distinctive antibody populations.

The nonimmunogenicity of linear 5 was not entirely expected since other linear peptides with V3 sequences are immunogenic (*23*, *24*, *27*). For example, Vu et al. (*27*) reported that a long C4−V3 MN peptide (39-residues; 23 from V3) generated high neutralizing titers to MN in rhesus monkeys (2000 to >5000) which showed not only that the C4−V3 MN peptide is immunogenic but implies that it was more immunogenic than loop 5. However, comparisons are complicated for instance by differences in neutralization assays. The neutralization sensitivity of HIV-1 can vary widely for ostensibly the same assay (*68*, *69*). Also, longer peptides might generate neutralizing antibodies to epitopes other than the SND. MAb 58.2 served as a reference MAb in this study indicating that loop 5 generated potent neutralizing antibodies to the SND.

The difference in the immunogenicity of linear 5 and loop 5 is structure-based and indicates that a degree of conformational compatibility must be achieved for antibodies to bind the SND on MN. Structure-based rationales have also been proposed to account for the immunogenicity of long V3 peptides (*11*, *27*, *56*). Various NMR studies show that long V3 peptides stabilize structures at various positions in aqueous solutions at 5 °C (*27*, *56*, *77*). It has been suggested from models based on NMR data that the C4−V3 MN peptide stabilizes a contiguous apolar surface at I1−H2−I3−G4−P5 (our numbering) and noted that this structure correlates with neutralizing activity against MN (*27*). Long V3 peptides also stabilize β-turns at GPGR (*11*, *27*, *56*, *77*) that have been proposed to account for V3 peptide immunogenicity (*11*, *56*). Evidence for a β-turn at GPGR was dependent on the length of V3 peptides (*56*, *77*). Many factors including peptide length, copeptide (*24*, *27*), carbohydrate (*77*), or conjugate protein (*78*) could alter the balance between disorder and structure that in turn may determine V3 peptide immunogenicity.

The immunogenicity of loop 5 could be due to its general loop shape and/or the stabilization of β-turns at GPGR and R(Aib)F(D-Ala). This compares with observations that MAbs 58.2, 50.1, and 59.1 bind β-turns at these positions and also neutralize MN. However, the types of β-turns and their relative orientations differ for the MAb bound conformers (Table 2) (*21*). These differences occur largely at GPGR suggesting that GPGR is flexible enough on MN to accommodate these differences. This apparent flexibility could also account for why P4 rgp120-specific antibodies and MAb 58.2 neutralize MN with similar potencies yet show wide differences in their relative affinities for loop 3 (Figure 3). Loop 3 is less flexible at PG due to the G6/D-Ala substitution. Consequently, it could be used to distinguish differences in the conformational preferences of MAb 58.2 and the P4 antibodies for this position (Figure 3). PG on MN is presumably more flexible than P(D-Ala) on loop 3 since both the P4 antibodies and MAb 58.2 neutralize MN with similar potencies although they presumably bind PG on MN in different conformations.

It can be questioned as to whether the MAbs 58.2, 50.1, 59.1, and the P4 antibodies complement the native SND as well as they might. Although the V3 MAbs were selected for neutralizing potency, they are technically antipeptide antibodies that in general are not anticipated to faithfully mirror native protein structure. If this were the case, then one might expect to find human V3 MAbs generated to

native conformations that fit MN better and are more potent. However, MAb 50.1 (and by extrapolation, the similarly potent MAb 58.2 and P4 antibodies) was found to be more potent against MN (IC_{50} = 0.08 μg/mL) than many V3 human MAbs (IC_{50} = 0.05−5 μg/mL) in the same assay (*39*). One of these HuMAbs, MAb 391/95-D (*29*, *39*), binds the same MN sequence as MAb 58.2 (*42*). Thus the data suggest that GPGR can readily undergo conformational transition(s) on MN.

Overall, this study demonstrates the enormous effect that conformational restriction can have on the antigenicity and immunogenicity of a short, synthetic peptide. It shows the stability of the hydrazone link under physiological conditions and the effectiveness of the loop-MAPS at generating native protein reactive antibodies. It also makes evident the simplifying role that an antibody template can play in identifying and improving a constrained peptide immunogen using enhanced affinity as a measure of fit. The hydrazone link could find wider use since it can constrain peptides to α-helices (*19*) as well as loops that are among the more commonly found substructures on protein surfaces. The effectiveness of an HIV-1 vaccine will likely depend on the degree to which it can generate immune responses to neutralizing epitopes defined by constant amino acids while excluding dependence on variable amino acids. Constrained peptides provide a means for focusing antibodies on defined epitopes. A continuing use of constrained peptides can be expected to lead to improved strategies and further perspectives on synthetic and HIV-1 vaccine research.

ACKNOWLEDGMENT

We thank Dr. Al Profy for providing MAbs 58.2 and 59.1, Dr. Lin-chiang Chiang at NuMega Resonance Labs, San Diego, for NMR data collection, and Dr. James Tam for stimulating discussions.

SUPPORTING INFORMATION AVAILABLE

Tables of NMR chemical shift data for the linear V3 Aib-peptide and loop 5. HPLC chromatograms for the purification of loop 5 and loop 5-MAPS. Mass spectra of MAPS-core, linear 5-MAPS, loop 5, and loop 5-MAPS. Titration curves of linear 5 and loop 5 antisera for rgp120. Competition ELISAs indicated in the text. Plots for temperature coefficients for the linear V3 Aib-peptide and loop 7. Portions of 600 MHz 2D ROESY spectra for loops 5 and 7 at pH 5.0, 10 °C. This information is available free of charge via the Internet at http:/pubs.acs.org.

REFERENCES

1. Burton, D. R., and Moore, J. P. (1998) *Nat. Med. 4 (5)*, 495−498.
2. Parren, P. W. H. I., Moore, J. P., Burton, D. R., and Sattentau, Q. J. (1999) *AIDS 13*, S137−S162.
3. *Human Retroviruses and AIDS 1998: Compilation and Analysis of Nucleic Acid and Amino Acid Sequences.* (1998) (Korber, B., Kuiken, C. L., Foley, B., Hahn, B., McCutchan, F., Mellors, J. W., and Sodroski, J., Eds.) Los Alamos National Laboratory, Los Alamos, NM.
4. Pilgrim, A. K., Pantaleo, G., Cohen, O. J., Fink, L. M., Zhou, J. Y., Zhou, J. T., Bolognesi, D. P., Fauci, A. S., and Montefiori, D. C. (1997) *J. Infect. Dis. 176*, 924−932.
5. Zhang, Y-J., Fracasso, C., Fiore, J. R., Björndal, A., Angarano, G., Gringeri, A., and Fenyö, E. M. (1997) *J. Infect. Dis. 176*, 1180−1187.
6. Van Regenmortel, M. H. V. (1996) *Methods*, *9*, 465−472.
7. Sela, M., and Arnon, R. (1992) *Vaccine 10*, 991−999.
8. Sela, M., Schechter, B., Schechter, I., and Borek, F. (1967) *Cold Springs Harbor Symposium on Quantitative Biology 32*, 537−545.
9. Jemmerson, R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 9180−9184.
10. Jemmerson, R., and Blankenfeld, R. (1989) *Mol. Immunol. 26*, 301−307.
11. Dyson, H. J., and Wright, P. E. (1995) *FASEB J. 9*, 37−42.
12. Sachs, D. H., Schechter, A. N., Eastlake, A., and Anfinsen C. B. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3790−3794.
13. Stanfield, R. L., Fieser, T. M., Lerner, R. A., and Wilson, I. A. (1990) *Science 248*, 712−719.
14. Muller, S., Plaue, S., Samama, J. P., Valette, M., Briand, J. P., and Van Regenmortel, M. H. V. (1990) *Vaccine 8*, 308−314.
15. Spangler, B. D. (1991) *J. Immunol. 146*, 1591−1595.
16. Arnon, R., Maron, E., Sela, M., and Anfinsen, C. B. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1450−1455.
17. Satterthwait, A. C., Arrhenius, T., Hagopian, R. A., Zavala, F., Nussenzweig, V., and Lerner, R. A. (1989) *Philos. Trans. R. Soc. London B 323*, 565−572.
18. Kaumaya, P. T., VanBuskirk, A. M., Goldberg, E., and Pierce, S. K. (1992) *J. Biol. Chem. 267*, 6338−6345.
19. Cabezas, E., and Satterthwait, A. C. (1999) *J. Am. Chem. Soc. 121*, 3862−3875.
20. Baker, E. N., and Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.
21. Stanfield, R. L., Cabezas, E., Satterthwait, A. C., Stura, E. A., Profy, A. T., and Wilson, I. A. (1999) *Structure 7*, 131−142.
22. LaRosa, G. J., Davide, J. P., Weinhold, K., Waterbury, J. A., Profy, A. T., Lewis, J. A., Langlois, A. J., Dreesman, G. R., Boswell, R. N., Shadduck, P., Holley, L. H., Karplus, M., Bolognesi, D. P., Matthews, T. J., Emini, E. A., Putney, S. D. (1990) *Science 249*, 932−935.
23. Nardalli, B., Lu, Y. A., Shiu, D. R., Delpierre-Defoort, C., Profy, A. T., and Tam, J. P. (1992) *J. Immunol. 148*, 914−920.
24. Ahlers, J. D., Pendleton, C. D., Dunlop, N., Minassian, A., Nara, P. L., and Berzofsky, J. A. (1993) *J. Immunol. 150*, 5647−5665.
25. Tolman, R. L., Bednarek, M. A., Johnson, B. A., Leanza, W. J., Marburg, S., Underwood, D. J., Emini, E. A., and Conley, A. J. (1993) *Int. J. Pept. Protein Res. 41*, 455−466.
26. Long, R. D., and Moeller, K. D. (1997) *J. Am. Chem. Soc. 119*, 12394−12395.
27. Vu, H. M., Myers, D., de Lorimier, R., Matthews, T. J., Moody, M. A., Heinly, C., Torres, J. V., Haynes, B. F., and Spicer, L. (1999) *J. Virol. 73*, 746−750.
28. Barbas III, C. F., Collet, T. A., Amberg, W., Roben, P., Binley, J. M., Hoekstra, D., Cababa, D., Jones, T. M., Williamson, R. A., Pilkington, G. R., Haigwoods, N. L., Cabezas, E., Satterthwait, A. C., Sanz, I., and Burton, D. R. (1993) *J. Mol. Biol. 230*, 812−823.
29. Gorny, M. K., Xu, J. Y., Karwowska, S., Buchbinder, A., and Zolla-Pazner, S. (1993) *J. Immunol. 150*, 635−643.
30. White-Scharf, M. E., Potts, B. J., Smith, L. M., Sokolowski, K. A., Rusche, J. R., and Silver, S. (1993) *Virol. 192*, 197−206.
31. King, D. S., Fields, C. G., and Fields, G. B. (1990) *Int. J. Pept. Protein Res. 36*, 255−266.
32. Carpino, L. A., Cohen, B. J., Stephens, Jr., K. E., Sadat-Aalaee, S. Y., Tien, J. H., and Langridge, D. C. (1986) *J. Org. Chem. Soc. 51*, 3732−3734.
33. Tam, J. P., Wu, C.-R., Liu, W., and Zhang, J.-W. (1991) *J. Am. Chem. Soc. 113*, 6657−6662.
34. Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70−77.
35. Tam, J. P. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5409−5413.

36. Men, Y., Gander, B., Merkle, H. P., and Corradin, G. (1996) *Vaccine 14*, 1442−1450.
37. Wu, J. X., and Satterthwait, A. C., unpublished.
38. Parren, P. W. H. I., Wang, M., Trkola, A., Binley, J. M., Purtscher, M., Katinger, H., Moore, J. P. and Burton, D. R. (1998) *J. Virol. 72*, 10270−10274.
39. VanCott, T. C., Bethke, F. R., Polonis, V. R., Gorny, M. K., Zolla-Pazner, S., Redfield, R. R., and Birx, D. L. (1994) *J. Immunol. 153*, 449−459.
40. Jellis, C. L., Cradick, T. J., Rennert, P., Salinas, P., Boyd, J., Amirault, T., and Gray, G. S. (1993) *Gene 137*, 63−68.
41. Moore, J. P., McCutchan, F. E., Poon, S. W., Mascola, J., Liu, J., Cao, Y., and Ho, D. D. (1994) *J. Virol. 68*, 8350−8364.
42. Seligman, S. J., Binley, J. M., Gorny, M. K., Burton, D. R., Zolla-Pazner, S., and Sokolowski, K. A. (1996) *Mol. Immunol. 33*, 737−745.
43. Rini, J. M., Stanfield, R. L., Stura, E. A., Salinas, P. A., Profy, A. T., and Wilson, I. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6325−6329.
44. Ghiara, J. B., Stura, E. A., Stanfield, R. L., Profy, A. T., and Wilson, I. A. (1994) *Science 264*, 82−85.
45. Ghiara, J. B., Ferguson, D. C., Satterthwait, A. C., Dyson, H. J., and Wilson, I. A. (1997) *J. Mol. Biol. 266*, 31−39.
46. Tugarinov, V., Zvi, A., Levy, R., and Anglister, J. (1999) *Nat. Struct. Biol. 6 (4)*, 331−335.
47. Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) *J. Mol. Biol. 206*, 759−777.
48. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1986) *Adv. Prot. Chem. 37*, 1−109.
49. Prasad, B. V. V., and Balaram, P. (1984) *CRC Crit. Rev. Biochem. 16 (4)*, 307−348.
50. Pease, L. G., Niu, C. H., and Zimmerman, G. (1979) *J. Am. Chem. Soc. 101*, 184−191.
51. Richardson, J. S., and Richardson, D. C. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., Ed.) pp 1−98, Plenum Press, New York.
52. Sarma, A. V. S., Raju, T. V., Kunwar, A. C. (1997) *J. Biochem. Biophys. Methods 34*, 83−98.
53. Wüthrich, K., Billeter, M., and Braun, W. (1984) *J. Mol. Biol. 180*, 715−740.
54. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
55. Dyson, H. J., and Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519−538.
56. Chandrasekhar, K., Profy, A. T., and Dyson, H. J. (1991) *Biochemistry 30*, 9187−9194.
57. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988) *J. Mol. Biol. 201*, 161−200.
58. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp 162−166, John Wiley, New York.
59. Wilmot, C. M., and Thornton, J. M. (1988) *J. Mol. Biol. 203*, 221−232.
60. Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
61. Wilmot, C. M., and Thornton, J. M. (1990) *Protein Eng. 3*, 479−493.
62. Basus, V. J. (1989) *Methods Enzymol. 177*, 132−149.
63. Defoort, J. P., Nardelli, B., Huang, W., and Tam, J. P. (1992) *Int. J. Pept. Protein Res. 40*, 214−221.
64. Panina-Bordignon, P., Tan, A., Termijtelen, A., Demotz, S., Corradin, G., and Lanzavecchia, A. (1989) *Eur. J. Immunol. 19*, 2237−2242.
65. DiGeorge, A. M., Wang, B., Kobs-Conrad, S. F., and Kaumaya, P. T. P. (1994) in *Peptides: Chemistry, Structure and Biology. Proceedings of the 13th American Peptide Symposium* (Hodges, R., and Smith, J. A., Eds.) pp 732−733, ESCOM Leiden, The Netherlands.
66. Jennings, V. M. (1995) *ILAR J. 37*, 119−125.
67. Hanly, W. C., Artwhol, J. E., and Bennett, B. T. (1995) *ILAR J. 37*, 93−118.
68. Layne, S. P., Merges, M. J., Spouge, J. L., Dembo, M., Nara, P. L. (1991) *J. Virol. 65*, 3293−3300.
69. D'Souza, M. P., Livnat, D., Bradac, J. A., Bridges, S. H. (1997) *J. Infect. Dis. 175*, 1056−1062.
70. Chesebro, B., Wehrly, K., Nishio, J., and Perryman, S. (1996) *J. Virol. 70*, 9055−9059.
71. Bou-Habib, D. C., Roderiquez, G., Oravecz, T., Berman, P. W., Lusso, P., and Norcross, M. A. (1994) *J. Virol. 68*, 6006−6013.
72. Teicher, E., Maron, E., and Arnon, R. (1973) *Immunochemistry 10*, 265−271.
73. Lerner, R. A. (1984) *Adv. Immunol. 36*, 1−44.
74. Calvo, J. C., Perkins, M., and Satterthwait, A. C. (1994) *Peptides: Chemistry, Structure and Biology; Proceedings of the Thirteenth American Peptide Symposium* (Hodges, R., Smith J. A., Eds.) pp 725−726, ESCOM, Leiden, The Netherlands.
75. Calvo, J. C., Chocontá, K. C., Diaz, D., Orozco, O., Espejo, F., Guzman, F., Patarroyo, M. E. (1998) *Peptides: Frontiers of Peptide Science; Proceedings of the 15th American Peptide Symposium* (Tam, J. P., and Kaumaya, P. T. P., Eds.) pp 819−820, Kluwer Academic, Dordrecht, The Netherlands.
76. Smith, A. D., Geisler, S. C., Chen, A. A., Resnick, D. A., Roy, B. M., Lewi, P. J., Arnold, E., and Arnold, G. F. (1998) *J. Virol. 72*, 651−659.
77. Huang, X., Barchi, J. J., Lung, F. D., Roller, P. P., Nara, P. L., Muschik, J., and Garrity, R. R. (1997) *Biochemistry 36*, 10846−10856.
78. Jelinek, R., Terry, T. D., Gesell, J. J., Malik, P., Perham, R. N., and Opella, S. J. (1997) *J. Mol. Biol. 266*, 649−655.
79. Parren, P. W., Mondor, I., Naniche, D., Ditzel, H. J., Klasse, P. J., Burton, D. R., and Sattentau, Q. J. (1998) *J. Virol. 72*, 3512−3519.

BI0003691

Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing Fabs

RL Stanfield[1], E Cabezas[1†], AC Satterthwait[1†], EA Stura[1‡], AT Profy[2§] and IA Wilson[1,3*]

Background: The third hypervariable (V3) loop of HIV-1 gp120 has been termed the principal neutralizing determinant (PND) of the virus and is involved in many aspects of virus infectivity. The V3 loop is required for viral entry into the cell via membrane fusion and is believed to interact with cell surface chemokine receptors on T cells and macrophages. Sequence changes in V3 can affect chemokine receptor usage, and can, therefore, modulate which types of cells are infected. Antibodies raised against peptides with V3 sequences can neutralize laboratory-adapted strains of the virus and inhibit syncytia formation. Fab fragments of these neutralizing antibodies in complex with V3 loop peptides have been studied by X-ray crystallography to determine the conformation of the V3 loop.

Results: We have determined three crystal structures of Fab 58.2, a broadly neutralizing antibody, in complex with one linear and two cyclic peptides the amino acid sequence of which comes from the MN isolate of the gp120 V3 loop. Although the peptide conformations are very similar for the linear and cyclic forms, they differ from that seen for the identical peptide bound to a different broadly neutralizing antibody, Fab 59.1, and for a similar peptide bound to the MN-specific Fab 50.1. The conformational difference in the peptide is localized around residues Gly-Pro-Gly-Arg, which are highly conserved in different HIV-1 isolates and are predicted to adopt a type II β turn.

Conclusions: The V3 loop can adopt at least two different conformations for the highly conserved Gly-Pro-Gly-Arg sequence at the tip of the loop. Thus, the HIV-1 V3 loop has some inherent conformational flexibility that may relate to its biological function.

Addresses: [1]The Scripps Research Institute, Department of Molecular Biology, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA, [2]Repligen Corporation, 117 Fourth Avenue, Needham, MA 02194, USA and [3]The Scripps Research Institute, Skaggs Institute for Chemical Biology, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.

Present addresses: [†]The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA, [‡]Department d'Ingenierie et d'Etude des Proteines, BAT 152 CEA/SACLAY, 91191 Gif sur Yvette Cedex, France and [§]Procept, Inc., 840 Memorial Drive, Cambridge, MA 02139, USA.

*Corresponding author.
E-mail: wilson@scripps.edu

Key words: antibody–peptide complex, Fab fragment, HIV-1, vaccine, X-ray crystallography

Received: **8 October 1998**
Revisions requested: **11 November 1998**
Revisions received: **23 November 1998**
Accepted: **1 December 1998**

Published: **29 January 1999**

Structure February 1999, **7**:131–142
http://biomednet.com/elecref/0969212600700131

 ISSN 0969-2126

Introduction

Human immunodeficiency virus 1 (HIV-1) is a member of the lentivirus subfamily of retroviruses. The exterior of the membrane-enveloped virus is embedded with multiple copies of the proteins gp120 and gp41, which are synthesized as a single protein precursor (gp160) and then cleaved, but remain in a noncovalent association on the membrane surface [1]. These viral antigens have been implicated in viral cell entry; gp120 binds to CD4 and then to at least one other coreceptor on the cell surface, before gp41-mediated fusion of the viral and target cell membranes. Recently, several β-chemokine receptors have been implicated as the viral secondary receptors, with CXCR4 serving as the primary coreceptor for T-cell tropic (T-tropic) or T-cell line adapted (TCLA) syncytia-inducing (SI) isolates and CCR5 serving as the major coreceptor for macrophage-tropic (M-tropic) nonsyncytia-inducing (NSI) isolates (for reviews, see [2,3]). Viruses that require CXCR4 for their coreceptors are termed 'X4' viruses, whereas viruses that require CCR5 are termed 'R5' viruses. Dual tropic viral strains have been identified that can use both coreceptors and, thus, are termed R5X4 viruses; some alternative coreceptors have also been identified that can be used by the virus to enter cells.

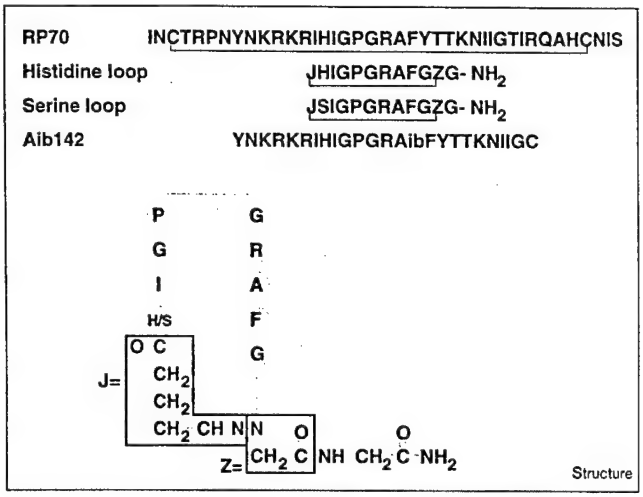
The third hypervariable (V3) domain of gp120 is a disulfide-linked loop of approximately 40 amino acids with a high degree of sequence diversity among different viral isolates [4–6]. The V3 loop is one of the major immunogenic sites on the virus and is sometimes called the principal neutralizing determinant (PND) [7]. The accessibility or exposure of the V3 loop on gp120 appears to vary depending on viral isolate type and increases significantly when the virus interacts with CD4 through a conformational change that is triggered in gp120 [8,9]. The V3 loop then becomes more sensitive to neutralizing antibodies and proteases as a result of the gp120–CD4 interaction. Sequence changes in V3 can alter viral cell tropism [10,11], antibody neutralization [12–14], neutralization of soluble CD4 [15], syncytium formation [11,16] and chemokine receptor usage [17]. It has

been noted that T-tropic V3 sequences are usually more basic in charge than M-tropic sequences through accumulation of positively charged residues in regions on either side of the GPGR tip of the loop [18,19]. As no easily discernable pattern relates specific changes in V3 residues to phenotypic changes, it is likely that conformational variations in gp120 and the V3 loop that arise from particular V3 loop sequences could modulate gp120 function. Although the crystal structure of a ternary complex between gp120, CD4 and an antibody Fab fragment has at last been determined [20–22], the gp120 was modified by removing several loops including V3. Thus, the paucity of structural information on the V3 loop still needs to be addressed.

In early work by La Rosa *et al.* [4–6], an analysis of 245 different HIV-1 V3 loop sequences showed that some amino acid positions in the loop are highly variable in amino acid composition, whereas others, especially those near the tip of the loop, are highly conserved; for example, examination of the MN sequence IHIGPGRAFY reveals that these amino acids are found at their respective positions in 94, 47, 82, 98, 95, 98, 91, 83, 72 and 80% of the other 244 isolates. The GPGR sequence is highly conserved in all known sequences of subtype B, whereas other group M isolates are characterized by the highly related GPGQ sequence at this position. The high degree of conservation of some of these amino acid positions suggests that they have a key structural and functional role and are important in some aspect of virus infectivity, such as modulation of the conformation of the V3 loop, or for modulating the interaction with coreceptor. The LaRosa *et al.* study [4–6] also predicted the secondary structure for their V3 consensus sequence (CTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHC) to consist of an extended β strand, a type II β turn around residues GPGR, a β strand and then an α helix. This secondary structural prediction was carried out for all 245 sequences, with 93% predicted to have the extended β strand around residues SIHI and 66% predicted to have the short helix at residues DIRQ. Pro, Gly and Arg are the most commonly found residues in positions i+1, i+2 and i+3 of a type II β turn, whereas Gln is the second most commonly found residue at position i+3 [23].

We have been investigating several HIV-1 neutralizing antibodies and their complexes with V3 peptides in order to determine the tertiary conformation of the V3 loop and to understand why some antibodies are viral-specific while others recognize many different viral strains. These neutralizing antibodies recognize the V3 epitope as a free peptide in the soluble gp120 protein and on the intact HIV-1 virus. Knowledge of the conformation of this loop may help to explain coreceptor usage and the changes that take place in the virus upon conversion from a primary M-tropic isolate to the T-tropic strains associated with disease progression. A panel of antibodies, which were raised against the same 40 amino acid V3 peptide RP70 [24] (Figure 1), neutralize the

Figure 1



Sequences of the V3 loop peptides used in crystallization of the peptide–Fab complexes. Fabs 58.2, 50.1 and 59.1 were all raised against the 40 amino acid, disulfide-linked peptide RP70, which has the MN viral sequence. A disulfide bond links the two cysteine residues as shown. Aib142 is the linear peptide crystallized with both 58.2 and 59.1. The Aib residue in peptide Aib142 is an α-aminoisobutyric acid [NHC(CH_3)$_2$C(O)] and replaces an Ala residue in the original sequence. The His loop and Ser loop cyclic peptides are constrained with a hydrazone linkage between residues J and Z, as shown in detail at the bottom of the figure. In the crystal structure of these cyclic peptides, the first residue (J) is referred to as Arn^{P1}, and consists of atoms N-CH-CH_2-CH_2-CH_2-C-O. The second residue is then His/Ser^{P315}, the third residue is Ile^{P316}, and so on. The last naturally occurring residue is Phe^{P324} which is then followed by Gly^{P10}, Gly^{P11}(Z) and Gly^{P12}. The N atoms of Arn^{P1} and Gly^{P11} are connected by the hydrazone linkage.

virus by preventing viral–cell membrane fusion. It has long been thought that these and other V3-specific antibodies do not interfere with CD4 binding. Recent studies using intact viral particles rather than soluble monomeric gp120, however, indicate that V3 neutralizing antibodies may also prevent the binding of HIV-1 to $CD4^+$ human cells [25].

Previously, we have reported the structures of two antibody Fab fragments, 50.1 and 59.1, in complex with V3 peptides [26–29]. Fab 50.1 is reported to be highly specific for the MN viral strain (T-tropic, SI), whereas 59.1 is reported to strongly neutralize IIIB (T-tropic, SI) and weakly neutralize MN [24]. The X-ray structures for V3 peptides in complex with Fabs 50.1 and 59.1 agreed well with the La Rosa *et al.* [4–6] predictions up to the conserved GPGR tip of the loop. In the 50.1–peptide complex, eight amino acid residues were ordered in the antibody combining site (CKRIHIGPG). Residues CKRIHI were in an extended β conformation with the two Ile residues buried in tightly fitting hydrophobic pockets in the antibody combining site and the GPG forming the start of a turn. The first Gly residue had torsion angles in the

epsilon (ε) region of Ramachandran space, a region usually only available to Gly residues, centered around $\phi = 90°$ and $\psi = 180°$. The PG had torsion angles appropriate for a type II β turn. In the 59.1–peptide complex, 10 amino acid residues were visible (HIGPGRAFYT). Five of these residues had structural counterparts in the peptide from the 50.1 complex (HIGPG), and displayed very similar mainchain torsion angles. Following the type II β turn around GPGR, we observed a double bend consisting of a type I turn for residues GRAF (in [28], this turn is referred to as type III; the type III turn is a variant of the type I turn and is usually classified as type I) and a type I turn around RAFY. The sidechain of the Arg residue is bound in a deep, negatively charged pocket on the antibody surface.

The individual peptide structures from the 50.1 and 59.1 complexes could be pieced or spliced together to make a more complete model for a 13-residue section of the V3 loop [29] that included the GPGR tip. Contrary to the previous predictions [4–6] of a single tight turn around GPGR, a much broader double turn was found that consisted of a type II turn around GPGR and a double bend (type I, GRAF; type I, RAFY) around GRAFY. The Ala residue in the double bend had α-helical torsion angles of $\phi = -60°$, $\psi = -45°$, similar to one of the two preferred conformations for the conformationally restricted residue Aib [α-aminoisobutyryl residue $-NHC(CH_3)_2C(O)$], which normally adopts torsion angles of –57°,–47° or +57°,+47° [30,31]. We surmised that replacement of the Ala residue with an Aib residue could constrain the peptide and provide a more rigid scaffold for vaccine or drug design around this template. Nuclear magnetic resonance (NMR) studies of such an Aib-containing peptide confirmed that additional structure was indeed conferred on the peptide in aqueous solution [29]; the X-ray structure determination of Fab 59.1 with the Aib-containing peptide showed that the Aib residue made no significant changes to the peptide conformation, as expected [29]. The Aib residue was not necessary for binding or for crystallization, as Fab 59.1 could be crystallized isomorphously with peptides that were identical except for the Ala→Aib mutation. Another attempt to restrict conformation of the V3 peptides involved using a synthetic hydrogen bond mimic incorporating a hydrazone bond. These hydrogen bond mimics are designed to replace a CO–HN hydrogen bond within a β-hairpin structure, thus cyclizing the peptide into a constrained hairpin structure.

Here, we describe the crystal structure of Fab 58.2, a highly potent and broadly neutralizing antibody that has been reported to neutralize both T-tropic and M-tropic viral strains, as well as several strains that have been passaged only once or twice [24]. We report the structures for Fab 58.2 in complex with one linear Aib-containing peptide and two cyclic, constrained peptides.

Results and discussion

Structure determination and refinement

The crystal structures for Fab 58.2 in complex with one linear (Aib142; Figure 1) and two hydrazone-linked cyclic peptides (His loop, Ser loop; Figure 1) have been determined to 2.0 Å (Aib142) and 2.8 Å (His loop and Ser loop). Data collection and refinement statistics are summarized in Table 1. All structures were determined by the molecular replacement (MR) method using previously determined Fab structures (see Materials and methods section). The crystals of the two Fab–cyclic peptide complexes were isomorphous, as a result of only a single amino acid difference in their peptides (His315→Ser). All structures were refined with the X-PLOR simulated annealing refinement protocol [32], using the Engh and Huber parameter set [33] and a bulk-solvent correction. The Fab molecules are numbered according to standard convention [34] with light and heavy chain identifiers L and H, respectively. The peptides are numbered according to the BH10 isolate sequence [35] with a P chain identifier (R^{P313} I^{P314} H^{P315} I^{P316} G^{P319} P^{P320} G^{P321} R^{P322} A^{P323} F^{P324} Y^{P325}). Note that the sequential numbering system is interrupted due to a two amino acid insert, after P316, that occurs in the IIIB isolate. The Fab 58.2–Aib142 coordinates contain Fab residues L1–L212, H1–H230 and peptide residues P313–P325 (11 residues). The Fab 58.2–His-loop coordinates contain Fab residues L1–L212, H1–H230 and the cyclic peptide with the complete hydrazone linker (P1–P12; see Figure 1). The Fab 58.2–Ser-loop coordinates contain Fab residues L1–L211, H1–H230 and the cyclic peptide minus the hydrazone linker (P2–P12).

The Fab–Aib142 peptide complex structure (Figure 2a) was determined to 2.0 Å resolution. Analysis of the Ramachandran plot by PROCHECK [36] shows that 89% of the residues are in most favored positions with three residues in disallowed regions. These 'disallowed' residues include two residues in a disordered loop in the constant heavy chain (residues H128–H136), which almost always displays high displacement parameters in other antibody structure determinations [37] and residue Ala^{L51} ($\phi = 69°$, $\psi = -53°$), which is the i+2 residue in a distorted type I' turn in most Fab L2 complementary determining region (CDR) loops [38].

Two Fab–cyclic peptide structures (Figure 2b,c) were determined, both to 2.8 Å resolution. These structures differ at only one residue in the peptide (Figure 1). The R_{free} was not monitored during the Ser loop refinement as this particular structure was determined prior to the introduction of the R_{free} validator. Only one bound water was included in each model due to the moderate resolution. Ramachandran plot analysis of the His loop complex by PROCHECK [36] shows that 81% of the residues are in most favored positions, with three residues in disallowed regions (two residues from the H128–H136 loop and Ala^{L51}).

Table 1

X-ray diffraction data and refinement statistics for Fab 58.2 complexes.

Peptide	Aib142	His loop	Ser loop
Space group	C2	$P2_12_12$	$P2_12_12$
Unit cell (Å)	72.9, 71.9, 88.2 $\beta = 98.3°$	96.1, 114.7, 49.5	95.5, 115.5, 49.6
Number of crystals	3	1	1
Resolution (Å)	43.6–2.0	24.0–2.8	47.8–2.8
Last shell (Å)	2.05–2.0	3.02–2.8	2.98–2.8
Observations	183,103 (9505)	44,763 (4678)	47,266 (9337)
Unique reflections	30,460 (2014)	12,923 (2287)	13,951 (2259)
R_{merge} (%)/ R_{sym} (%)	10.6 (54.7)	7.6 (28.8)	7.2 (21.8)
Completeness (%)	99.6 (98.5)	91.9 (82.8)	98.9 (97.7)
$\langle I/\sigma_I \rangle$	29.5 (3.4)	7.5 (1.0)	13.2 (1.5)
Generator	Siemens M18X; Rigaku RU-300	Rigaku RU-200	Siemens M18X
Optics	Supper long mirrors; monochromator	Monochromator	Supper double mirrors
Detector	Mar 30 cm; R-axis II	Xuong-Hamlin multiwire	Siemens X1000
Data reduction program	HKL [81]	Xuong-Hamlin [82]	XENGEN [83]
R_{cryst} (%)*	19.6 (31.8)	18.5 (27.6)	20.0 (37.7)
R_{free} (%)†	25.6 (38.2)	30.5 (36.6)	–
Rmsd bond lengths (Å)	0.010	0.012	0.008
Rmsd bond angles (°)	1.82	1.87	1.59
Number of water molecules	120	1	1

Values given in parentheses are for the highest resolution shell. *$R_{cryst} = \Sigma_h \| \|F_{obs}(h)\| - \|F_c(h)\| \| / \Sigma_h \|F_{obs}(h)\|$, where F_{obs} and F_c are the observed and calculated structure-factor amplitudes for reflection $h = (h,k,l)$. †R_{free} is calculated with 10% (5% for the His loop–Fab complex) of the data never used in the refinement. R_{free} was not calculated for the Ser loop complex. The R_{cryst} and R_{free} are calculated using all data with $F/\sigma_F > 0$. A bulk-solvent correction as implemented in X-PLOR [72] was applied to all data sets.

The Ser loop complex has 87% of the residues in most favored regions with five residues in disallowed regions (two in the H128–H136 loop, the C-terminal residue of the heavy chain, Ala^{L51} and Lys^{H64}, which is the i+1 residue of a type II′ turn).

Structure description and comparisons

Fab 58.2 is a mouse IgG1 κ antibody. The elbow angles for Fab 58.2 are 174.3°, 144.8° and 144.3° in the Aib142, His loop and Ser loop complexes, respectively. The 58.2 CDR loops L1, L2, L3 and H2 belong to canonical classes 5, 1, 1 and 1, as classified previously [39,40]. The 58.2 H1 CDR loop has a one amino acid insert following residue 35 and so should belong to canonical class 2. Other Fabs with H1 loops of this length are AN02 (PDB accession code 1baf; [41]) and N10 (1nsn; [42]); however, the H1 loops in both 58.2 and N10 differ from the expected canonical structure around several residues at their tips and also differ in this region from each other (Figure 3). Not surprisingly, a tendency for an increase in flexibility is seen as the H1 loop increases in length. Fab 58.2 has a long CDR H3 with a nine amino acid insert after residue 100. Previously determined structures with H3 loops of this length include KOL, 3D6 and R45-45-11 (accession codes 2fb4 [43], 1dfb [44] and 1ikf [45], respectively). The conformation of the base of the H3 loop is dictated by residues Ala^{H93}, Arg^{H94} and Asp^{H101}. This combination of residues results in a 'kinked' base structure, as expected [39,46], with the sidechain of Asp^{H101} forming a salt bridge with Arg^{H94} and the mainchain carbonyl oxygen of H100I (the residue preceding H101) forming a hydrogen bond to Trp^{H103} Nε1.

The linear Aib142 peptide is 24 amino acids with an Aib residue replacing Ala^{P323} (Figure 1). Electron density is clearly visible for 11 of the 24 residues, RIHIGPGRAibFY (Figure 4a,b). The peptide has an extended conformation for residues RIHI, which connect to a type I ($\alpha_R \rightarrow \alpha_R$) turn around residues GPGR. A type VIa ($\beta_P \rightarrow \alpha_R$) turn is found for residues GRAibF and a type I turn is found around residues RAibFY, making residues GRAibFY a double bend. Residue Arg^{P322} is bound in a deep pocket in a highly negatively charged antibody combining site (Figure 5a), forming charge–charge interactions with residues Asp^{L94} and Glu^{H95} (Figure 5b). The bound portions of the cyclic peptides have a very similar conformation to the bound portions of the linear peptide. Clear electron density is present for most of the cyclic peptides, with a small break in density around the hydrazone linker in the His loop complex, and weak density for the linker in the Ser loop complex (Figure 4c,d). The hydrazone linker has been left out of the Ser loop model because of its weak electron density.

The three peptides bind to the antibody in virtually identical orientations, with the peptide lying flat on the Fab surface and residue Arg^{P322} protruding into a deep,

Figure 2

Stereoviews of the X-ray structures of the Fab 58.2–peptide complexes. **(a)** The Fab 58.2–Aib142 complex. The Cα trace is shown for the light chain (cyan) and heavy chain (blue) of the Fab. All atoms of the bound peptide are shown in red. Every tenth Cα atom is highlighted with a sphere and some atoms are labeled. **(b)** The Fab 58.2–His loop complex. **(c)** The Fab 58.2–Ser loop complex. Figures were made with the program MOLSCRIPT [74].

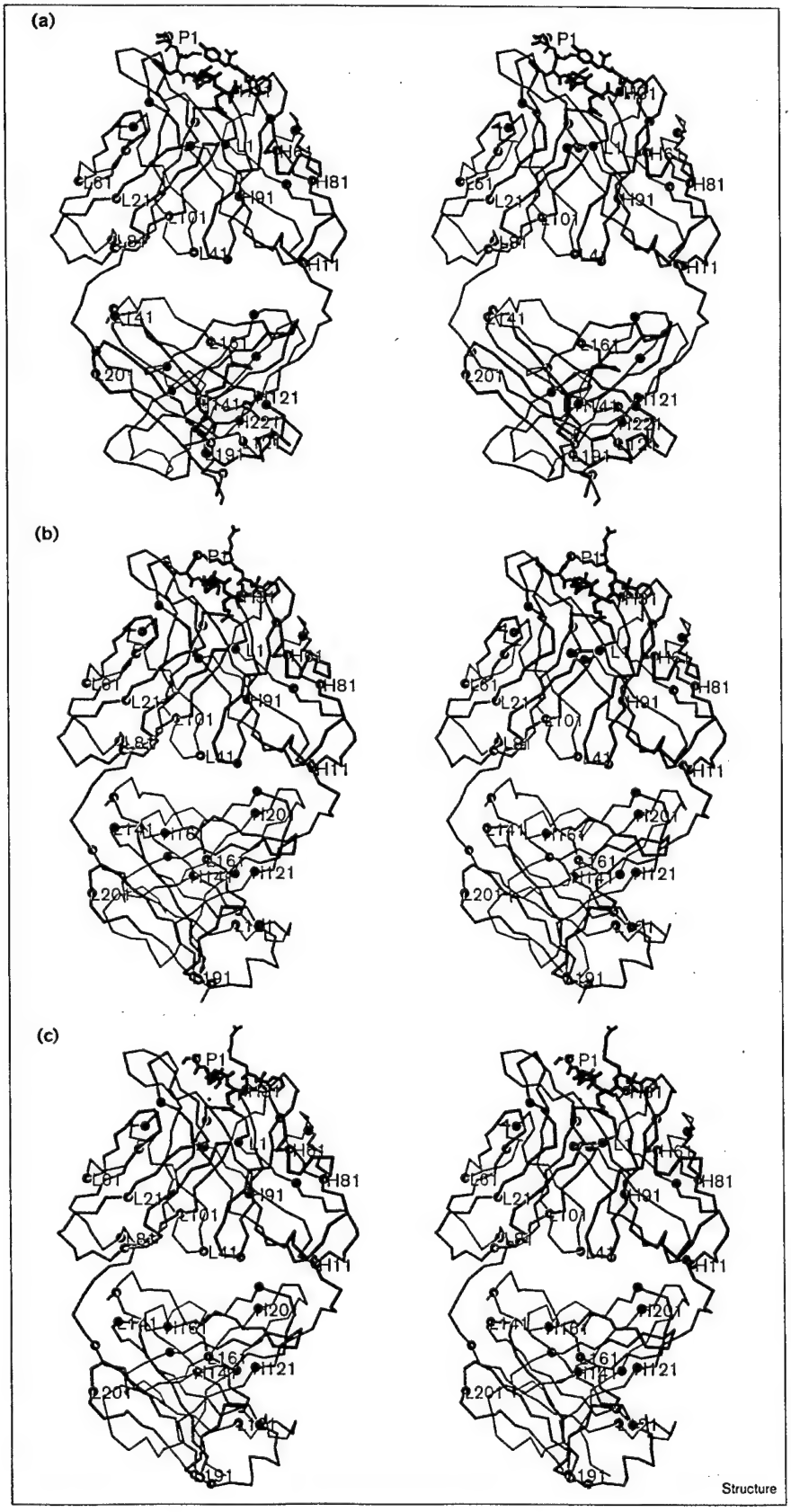
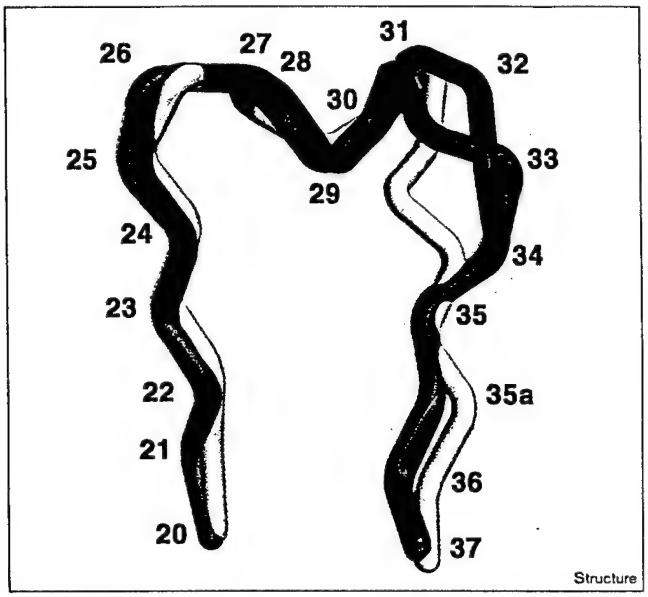


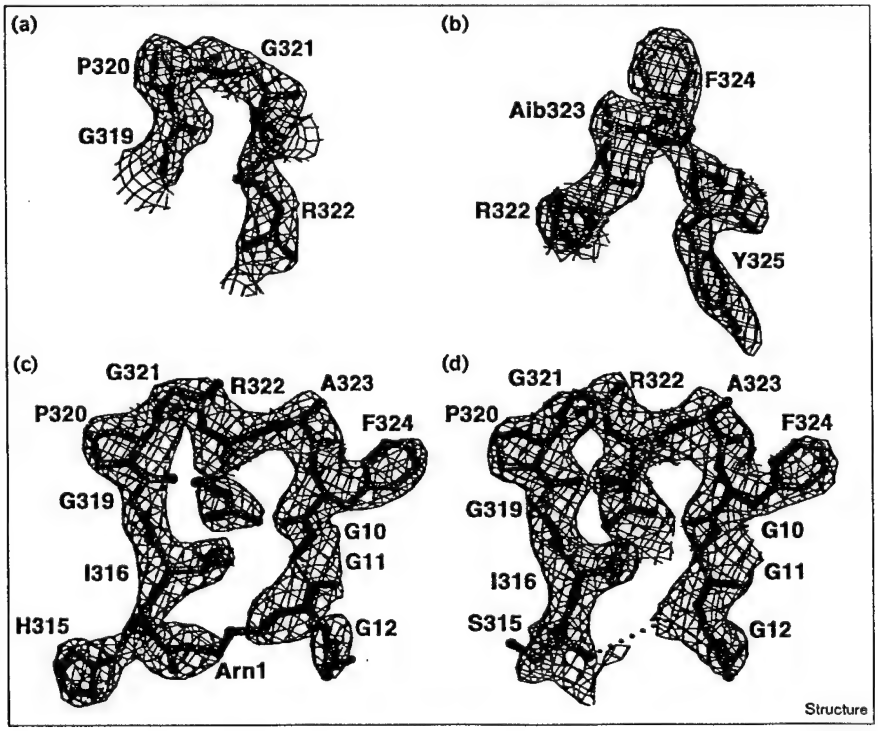
Figure 3



Comparison of the conformation of the H1 loops of Fabs 58.2 (red), N10 (yellow) and AN02 (blue). The coordinates of Fabs N10 and AN02 were obtained from the Protein Data Bank (accession codes 1nsn and 1baf, respectively).

negatively charged pocket on the Fab surface (Figure 5a). The surface area [47] buried by the peptide on the Fab is 666, 609 and 597 Å^2 for the Aib142, His loop and Ser loop peptides, respectively. The corresponding buried surface area on the peptides is 578, 540 and 487 Å^2. The Fab uses five of its six CDR loops to bind peptide; CDR L2 is not used and CDR H1 makes only minor contacts. This pattern of CDR usage is typical for antihapten and antipeptide antibodies [48]. A total of 124, 103 and 77 van der Waals contacts are made between the Fab and peptide for the Aib142, His loop and Ser loop complexes, respectively (Table 2). In each complex, charge–charge interactions are found between Arg^{P322} of the peptide and Asp^{L94} and Glu^{H95} of the Fab (Table 3). In the Aib142 peptide, an additional charge–charge interaction is made between Arg^{P313} and Asp^{L28}. A total of 11, 7 and 7 hydrogen bonds and salt bridge interactions are made between the peptides and Fab in the Aib142, His loop and Ser loop complexes, respectively (Table 3). The size and features of this interaction surface are similar to those of other Fab–peptide complex structures [49]. The Fab conformations are very similar between the different complexes as shown by the small differences in V_L–V_H pairings (less than 1.4°) between the complexes and the root mean square deviations (rmsds) for the individual domains, which range from 0.31 Å to 1.20 Å. One loop in the constant heavy chain (H128–H136), which frequently shows high thermal displacement parameters in Fab crystal structures [37], accounts for the high rmsds (0.83–1.20 Å) for the C_H1 regions. Removal of this loop from rmsd calculations results in rmsds of 0.61–0.72 Å.

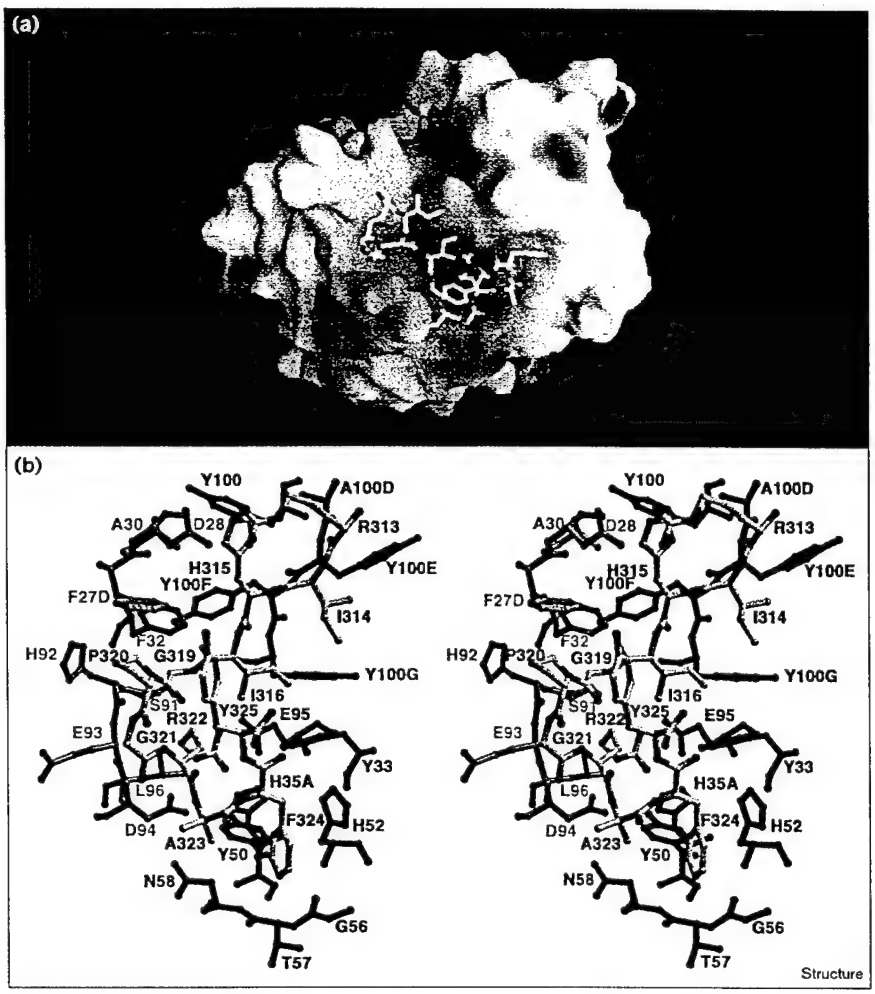
Figure 4



Electron density for the V3 peptides bound to Fab 58.2. **(a)** In the electron density for the 58.2–Aib142 complex, 11 residues of the 24-mer peptide Aib142 are visible. The simulated annealing (SA) omit electron-density map for peptide Aib142 is shown for the region around residues GPGR, which form a type I β turn. **(b)** SA omit electron density for peptide Aib142 showing the region around residues RAibFY. **(c)** SA omit electron density for the His loop peptide. **(d)** SA omit electron density for the Ser loop peptide. The hydrazone linker was not included in this model due to weak electron density; the approximate position of the linker is indicated by a dotted line. All maps are contoured at a 1σ level. Figures were made with BOBSCRIPT [75] and Raster3D [76,77].

Figure 5

The Fab 58.2 binding pocket with the Aib142 peptide bound. **(a)** GRASP [78] representation of the antibody combining site with Aib142 peptide bound. Negatively charged regions are colored in red and positively charged regions are colored in blue. **(b)** Stereoview of the combining site. The contacting residues of the light and heavy chains are shown in cyan and blue, respectively; the peptide is in yellow.



Although the resolution of the His loop and Ser loop structures is insufficient to allow accurate placement of ordered water molecules, one strong peak of density in the Fab binding site of all three structures was assigned as a water position. Additionally, 120 water molecules were included in the 2.0 Å resolution refinement of the Aib142 complex. Six of these water molecules are found buried at the Fab–peptide interface (W12, W41, W62, W67, W72, W84).

The conclusions from epitope mapping experiments of antibody 58.2 agree well with the structural results. The epitope specificity for 58.2 was determined previously by screening a peptide display library expressing 1.5×10^8 unique 20 amino acid peptide sequences against 58.2 [50], as well as by epitope mapping, where 266 unique 14-mer peptides were prepared, with each amino acid position replaced with all amino acids (except Cys) (peptide KRKR-IHIGPGRAFY) [50]. Residues KRKRI (P310–P314) could be replaced by most of the other amino acids, indicating that they are not important for peptide binding to Fab 58.2. His^{P315} could be replaced with any amino acid except Gly, Pro and Thr. Ile^{P316} could only be replaced with Glu, Leu, Met and Val. Gly^{P319} could not be replaced by any other amino acid. Pro^{P320} could only be replaced by Ile. The second Gly^{P321} was not so important, as it could be replaced with all amino acids except Ile, Pro, Ser, Thr, Val and Trp. Arg^{P322} could not be substituted, whereas Ala^{P323} could be replaced with any amino acid except Pro. The Phe^{P324} could be replaced with Asp, Ile, Lys, Gln, Arg, Ser, Thr, Val and Trp. These epitope mapping results indicate that residues Gly^{P319}, Pro^{P320} and Arg^{P322} are especially important for antibody binding, in agreement with the crystal structure that shows that Arg^{P322} forms a charge–charge interaction with the Fab and residues Gly^{P319} and Pro^{P320} are critical for determining the conformation of the tip of the V3 loop. Residue Gly^{P319}, which of course is a highly conformationally flexible residue, may also be important for allowing the loop to adjust to more than one distinct conformation. Gly^{P321}, the i+2 residue in the type I β turn of the peptide in the 58.2–peptide complexes, is not absolutely

Table 2

Van der Waals contacts between Fab 58.2 and bound peptides.

Peptide residue	In all complexes	Aib142	His loop	Ser loop
Arg^{P313}		Asp^{L28} Ala^{H100D} Tyr^{H100E}		
Ile^{P314}/Arn^{P1}		Tyr^{H100E} Tyr^{H100G}		
His^{P315}/Ser^{P315}	Tyr^{H100E}	Asp^{L28} Ala^{L30} Phe^{L32} Tyr^{H100} Ala^{H100D} Tyr^{H100F} Tyr^{H100G}	Asp^{L28} Ala^{L30} Phe^{L32} Tyr^{H100}	
Ile^{P316}	Phe^{L32} Tyr^{H33}	Tyr^{H100G}		Phe^{L27D} Tyr^{H100G}
Gly^{P319}	Phe^{L32} His^{L92}	Ser^{L91}	Ser^{L91}	
Pro^{P320}	Phe^{L27D} His^{L92}			
Gly^{P321}	His^{L92} Glu^{L93} Asp^{L94}			
Arg^{P322}	Ser^{L91} Glu^{L93} Asp^{L94} Leu^{L96} Glu^{H95}	His^{H35a} Tyr^{H50}		Tyr^{H50}
Ala^{P323}/Aib^{P323}	Asp^{L94}	Tyr^{H50} Asn^{H58}		Asn^{H58}
Phe^{P324}	Tyr^{H50} His^{H52} Gly^{H56} Thr^{H57}		Tyr^{H33} Ser^{H54}	Tyr^{H33}
Tyr^{P325}/Gly^{P10}		Phe^{L27D}	Tyr^{H33} His^{H52}	Tyr^{H33} His^{H52}
Gly^{P11}			Tyr^{H53}	

Van der Waals contacts were calculated with the program CONTACSYM [84] using standard van der Waals radii [85]. The Fab residues contacted in all complexes are shown first, followed by additional contacts that are specific to the individual complexes.

required, in agreement with the diversity of residues that are found at the i+2 position of type I turns [23]. This result contrasts with the alanine replacement experiment on peptide binding to Fab 59.1 [24,28]. When Gly^{P321} is replaced by an Ala residue, the affinity of Fab 59.1 for the peptide is reduced to about 32% of that of the native peptide, in good agreement for the strong preference for a Gly residue at the i+2 position of a type II turn [23], as found in the 59.1–peptide complex.

Two different V3 conformations

The most notable feature of the V3 peptide conformation while bound to Fab 58.2 is that it differs from the two previously determined structures for the same or similar peptide in complex with two other Fabs. The previously determined V3 loop structures were for peptide complexes with Fab 50.1, where the peptide epitope was identified as CKRIHIGPG, and with Fab 59.1, where the peptide epitope was HIGPGRAFYT. The three antibodies were all generated against the same peptide immunogen, RP70 [24] (Figure 1). The five residues (HIGPG) that are structurally ordered in peptides bound to both Fabs 50.1 and 59.1 could be superimposed with low rmsds (Figure 6) [28,29]. These two peptides both adopt torsion angles for a type II β turn around residues GPGR. Thus, it was surprising to observe that the peptide bound to Fab 58.2 had torsion angles for a type I β turn around these same residues. The mainchain torsion angles for these four residues (GPGR) differ significantly between the two peptide-conformational types. The residues before and after the GPGR sequence, however, share very similar mainchain torsion angles (Table 4). The second β turn, around residues GRAibF (or GRAF), is type I in the 59.1 complex (type I [$\alpha_R \rightarrow \alpha_R$] turns have average torsion angles

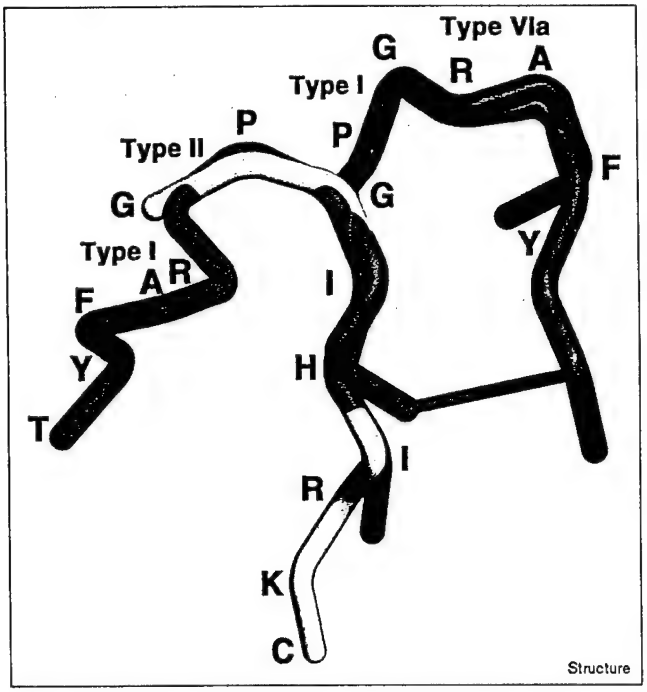
Table 3

Hydrogen bonds and salt bridge interactions (Å) in Fab 58.2–peptide complexes.

Peptide atom			Fab atom			Aib142	His loop	Ser loop
P313	Arg	NH2	L28	Asp	OD2	2.8		
P313	Arg	NH1	L28	Asp	O	3.2		
P313	Arg	O	H100e	Tyr	N	2.9		
P315	His/Ser	N	H100e	Tyr	O	2.7		
P315	His/Ser	ND1	H100e	Tyr	O	3.3		
P319	Gly	N	L92	His	O	3.0	2.9	3.1
P321	Gly	O	L94	Asp	N	3.3	3.3	3.3
P322	Arg	NE	L94	Asp	OD1	2.8		2.9
P322	Arg	NH2	H95	Glu	OE1	3.0	2.8	2.8
P322	Arg	NH1	H95	Glu	OE2	2.9	2.8	3.0
P323	Aib/Ala	N	L94	Asp	OD2	2.9	3.1	2.9
P324	Phe	O	H33	Tyr	OH		2.8	2.7
P10	Gly	O	H52	His	NE2		3.0	

Hydrogen bonds were calculated with HBPLUS [86].

Figure 6



Dual conformations of the V3 loop. The peptides bound to Fabs 50.1, 59.1 and 58.2 are superimposed using the Cα atoms of residues with similar torsion angles (Table 4). The 50.1 peptide is in yellow, the 59.1 peptide is in purple, and the Aib142 and His loop peptides bound to 58.2 are colored in blue and green, respectively. The hydrazone linkage in the His loop peptide is indicated by a straight connection. All four peptides share similar mainchain torsion angles for residues IHI and AF, the residues on either side of GPGR. The GPGR region in the 58.2 peptide (type I turn) differs from that seen in the 50.1 and 59.1 peptides (type II turn). Figure made with the program MIDAS [79,80].

Table 4

Mainchain dihedral angles (°) for V3 peptides bound to Fabs 50.1, 59.1 and 58.2.

	50.1 Peptide		59.1 Peptide		58.2 Aib142 peptide	
	ϕ	ψ	ϕ	ψ	ϕ	ψ
Cys^{P311}						
Lys^{P312}	−119	−169				
Arg^{P313}	−144	135				163
Ile^{P314}	−67	124			−122	144
His^{P315}	−89	134		144	−107	110
Ile^{P316}	−124	141	−133	−169	−124	137
Gly^{P319}	102	173	79	−172	−58	141
Pro^{P320}	−63	170	−73	165	−63	−24
Gly^{P321}	86		60	9	−95	−1
Arg^{P322}			−64	−41	−65	141
Ala/Aib^{P323}			−60	−46	−50	−42
Phe^{P324}			−94	19	−107	40
Tyr^{P325}			−49	−22	−68	161
Thr^{P326}			−126			

of ϕ = −60°, ψ = −30° and ϕ = −90°, ψ = 0° for the i+1 and i+2 residues), and type VIa in the 58.2 complex (type VIa [$\beta_P \rightarrow \alpha_R$] turns have average torsion angles of ϕ = −60°, ψ = 120° and ϕ = −90°, ψ = 0° for the i+1 and i+2 residues) with the torsional angles of residues AFY being highly similar in all peptide complexes (Table 4). So, the peptides bound to Fabs 59.1 and 58.2 both contain a β turn followed by a double bend, but the different torsion angles for the central GPGR residues change the precise character of these turns and bends. NMR studies of the Aib142 peptide in solution have provided evidence for populations of both type I and II turns around residues GPGR [29], which is in good agreement with the two peptide conformations that we observe in the crystal structures of the V3 loop with different Fabs.

Several other NMR studies of V3 peptide conformation have been carried out in solution for both linear and cyclic peptides of varying length and sequence. These studies have shown that the V3 loop by itself is largely disordered in water, but contains a population of tight turns for residues GPGR. The NMR studies show evidence for type II turns [51,52], both I and II turns [53–56], nonspecific β turns [57] and a double turn, similar to that seen in the 59.1-peptide X-ray structure [58]. NMR studies have also supported the proposal that glycosylation of the V3 peptide can affect its conformation [59,60] and that sequence changes affect the conformation [61]. NMR studies of an antibody–V3 loop peptide complex have also been carried out [62,63] but indicated only that the peptide forms a loop with residues GPGR at the tip, with no more detailed information being reported. Interestingly, a search of the Brookhaven Protein Data Bank [64] for proteins containing the sequence GPGR identified two different proteins, but neither were in a type I or II β turn. In scorpion toxin II (PDB codes 1ptx and 1aho [65]), these residues form a wide, undefined turn near the C terminus of the protein (i+1, ϕ = −65°, ψ = 150°; i+2, ϕ = 115°, ψ = 135°; $\beta_P \rightarrow \varepsilon$), whereas in chaperonin-10 (PDB code 1lep [66]) the GPGR residues are part of a β strand. Additionally, in a recently reported X-ray structure [67] of the major histocompatibility complex (MHC) class I H-$2D^d$–HIV-1 peptide (RGPGRAFVTI) complex, the GPGR is in an extended conformation, as expected for peptides complexed with MHC class I molecules.

Correlation of structure with function

We clearly observe two different conformations for the V3 loop while bound to different neutralizing antibodies, and two conformations for the tip of the loop in solution [29]. These V3 loop conformations are biologically relevant, as the antibodies bind to viral gp120 to effect their neutralization activity. These conformations may represent two different and distinct secondary structural motifs that may or may not relate to viral phenotype, or they could merely indicate that the V3 loop is flexible and able to

adopt different conformations. The only reported crystals of gp120 to date have been obtained for protein where this loop was truncated at its base [20–22]; crystals of gp120 with the intact loop have not yet been reported.

Biological implications

Human immunodeficiency virus 1 (HIV-1) is a membrane-enveloped virus, the exterior of which is embedded with multiple copies of the proteins gp120 and gp41. The third hypervariable loop (V3) of gp120 is involved in many aspects of virus infectivity and has, therefore, been the subject of much interest. The high degree of conservation of the Gly-Pro-Gly-Arg/Gln motif at the tip of the gp120 V3 loop, surrounded by regions of high sequence diversity, suggests this structural conservation is related to biological function.

Many studies have shown that the alteration of one or more amino acid residues within the stem of the V3 loop leads to changes in viral tropism. However, it remains to be seen whether these individual residue changes alter the loop structure or effect a switch in β-turn conformation at the tip of the loop. Clearly, a direct consequence of different turn types around Gly-Pro-Gly-Arg is that the flanking regions of V3 would be in different environments based on which turn type was found at the loop tip. The determination of the crystal structures of Fab fragments in complex with V3 loop peptides, presented here, goes some way towards understanding the conformational flexibility of this loop, although the answers to some questions remain elusive. These unanswered questions provide an important direction for future studies designed to reveal the biological role of the V3 loop. Further knowledge of the V3 structure will open up the possibilities for small molecule drug design based on its three-dimensional structure, or vaccine development through stabilization of biologically relevant conformations.

Materials and methods

The linear Aib142 peptide was prepared by chemical synthesis using Nα-9-fluorenylmethoxyl-carbonyl (Fmoc) protected amino acids and Fmoc-Rink's amide resin on an Advance Chemtech ACT350 peptide synthesizer, following the manufacturer's protocol. The N terminus was acetylated and the sidechain of Cys was protected with an acetamidomethyl (Acm) group. A mixture of trifluoroacetic acid (TFA)/phenol/ethanedithiol/thioanisole/water (84:6:2:4:4 v/w/v/v/v) was used to cleave the peptide from the resin. Peptide was purified using a preparative Cosmosil C-18 reverse-phase column on a Gilson high-performance liquid chromatography (HPLC) system. The peptide was eluted with a water/acetonitrile gradient containing 0.1% trifluoroacetic acid; 0–20% acetonitrile over 5 minutes followed by 20–35% acetonitrile over 20 min at 8 ml/min. The pure peptide eluted as a single peak on analytical HPLC and its identity was confirmed by mass spectrometry. The cyclic peptides were synthesized by solid phase synthesis using modified amino acids to form the hydrazone link. They were purified to single peaks by HPLC and their identities confirmed by mass spectrometry.

All crystallizations were carried out using the sitting drop, vapor diffusion method at 22.5°C. Crystals of the Fab 58.2–His loop and Ser loop complexes were obtained with a well solution containing 16% PEG 4000, 0.2 M imidazole malate, pH 5.0. The sitting drop consisted of 2.5 µl well solution, and 2.5 µl peptide/Fab (15 mg/ml Fab), with a peptide:Fab molar ratio of 2.4:1. Crystals of the Fab 58.2–Aib142 complex were grown from a well solution of 16% PEG 10,000, 0.2 M imidazole malate, pH 6.3. The drops consisted of 2.5 µl peptide/Fab (15 mg/ml Fab), and 2.5 µl well solution with a peptide:Fab molar ratio of 1.7:1. Space groups and unit cell dimensions for all crystal forms were determined by precession photography. Data collection statistics and information are tabulated for all three crystal forms in Table 1.

The Fab 58.2/Ser loop complex structure was determined first, using coordinates of Fab 59.1 (1acy, [28]), minus its peptide, as the MR model. Fabs 59.1 and 58.2 have approximately 83% and 65% sequence identity respectively in their V_L and V_H domains; their C_L and C_H1 domains are identical (IgG1, κ). The structure was determined using the X-PLOR rotation function and PC refinement [68], and a modified Harada translation function [69] written in our laboratory by D Filman and J Arévalo, with the intact 59.1 as a model. The intact Fab was found to be a suitable starting model because its elbow angle (the angle between the pseudo twofold axes of the $V_L - V_H$ and $C_L - C_H1$ domains) differed from that of the unknown Fab by only 11° (the elbow angle for 58.2 was 144° and for 59.1 was 135°). Alternating cycles of model building with the Alberta/Caltech version of TOM/FRODO [70,71] and refinement with the slow-cooling protocol in X-PLOR [72] were carried out, with peptide electron density being clearly visible after several cycles (Figure 4d). The refined coordinates for Fab 58.2–Ser loop (minus the peptide) were used as a starting model for the isomorphous Fab 58.2–His loop complex (Figure 4c). The Fab coordinates were refined using rigid body refinement followed by alternating cycles of slow-cooling refinement and model building as for the Ser loop complex. The Fab 58.2–Aib142 structure was also determined using MR. To estimate the elbow angle for the Fab 58.2–Aib142 structure, the Crowther Fast Rotation Function, as implemented in Merlot [73], was used to carry out cross-rotation functions with over 100 available Fab coordinate sets, which had previously been superimposed on a reference position, with the elbow angle aligned along the z-axis. A plot of elbow angle versus peak height for the top solution for each model suggested an elbow angle of around 180° for the unknown Fab 58.2. The starting Fab 58.2 model coordinates (the refined 58.2–Ser loop coordinates, without peptide) were artificially given this approximate elbow angle. The rotation function, PC refinement and translation function from X-PLOR were used for the structure determination [68]. The actual elbow angle for the refined structure was 175°. The structure was refined and rebuilt in alternating cycles as for the two previous complexes. Density for the peptide was very clear after only one cycle of slow-cooling refinement and manual rebuilding of the starting Fab model (Figure 4a,b).

Accession numbers

The coordinates and structure factors for the three Fab 58.2–peptide complexes are deposited in the PDB with accession codes 1f58, 2f58 and 3f58, to become available at the time of publication.

Acknowledgements

We thank Xiaoping Dai, Ron Hamlin, Rashid Syed and Tim Osslund for help with data collection and reduction for the Aib142–peptide complex crystals, and Ying Su and Nguyen-Huu Xuong for help with data collection and reduction for the Ser loop peptide complex. This work was supported by NIH grants GM-46192 (IAW) and AI-37512 (ACS), and carried out while RLS was a Scholar of the American Foundation for Aids Research. This is publication #11968-MB from The Scripps Research Institute.

References

1. McCune, J.M., *et al.*, & Weissman, I.L. (1988). Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell* **53**, 55-67.
2. Moore, J.P., Trkola, A. & Dragic, T. (1997). Co-receptors for HIV-1 entry. *Curr. Opin. Immunol.* **9**, 551-562.
3. Broder, C.C. & Collman, R.G. (1997). Chemokine receptors and HIV. *J. Leukoc. Biol.* **62**, 20-29.

4. LaRosa, G.J., *et al.*, & Putney, S.D. (1990). Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant. *Science* **249**, 932-935.
5. LaRosa, G.J., *et al.*, & Putney, S.D. (1991). Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant: further clarifications. *Science* **253**, 1146.
6. LaRosa, G.J., *et al.*, & Putney, S.D. (1991). Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant: corrections and clarifications. *Science* **251**, 811.
7. Javaherian, K., *et al.*, & Matthews, T.J. (1989). Principal neutralizing domain of the human immunodeficiency virus type 1 envelope protein. *Proc. Natl Acad. Sci. USA* **86**, 6768-6772.
8. McKeating, J.A., Cordell, J., Dean, C.J. & Balfe, P. (1992). Synergistic interaction between ligands binding to the CD4 binding site and V3 domain of human immunodeficiency virus type I gp120. *Virology* **191**, 732-742.
9. Sattentau, Q.J. & Moore, J.P. (1991). Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J. Exp. Med.* **174**, 407-415.
10. Hwang, S.S., Boyle, T.J., Lyerly, H.K. & Cullen, B.R. (1991). Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science* **253**, 71-74.
11. Chesebro, B., Wehrly, K., Nishio, J. & Perryman, S. (1996). Mapping of independent V3 envelope determinants of human immunodeficiency virus type 1 macrophage tropism and syncytium formation in lymphocytes. *J. Virol.* **70**, 9055-9059.
12. di Marzo Veronese, F., *et al.*, & Lusso, P. (1993). Loss of a neutralizing epitope by a spontaneous point mutation in the V3 loop of HIV-1 isolated from an infected laboratory worker. *J. Biol. Chem.* **268**, 25894-25901.
13. Wrin, T., Loh, T.P., Vennari, J.C., Schuitemaker, H. & Nunberg, J.H. (1995). Adaptation to persistent growth in the H9 cell line renders a primary isolate of human immunodeficiency virus type 1 sensitive to neutralization by vaccine sera. *J. Virol.* **69**, 39-48.
14. Ivanoff, L.A., *et al.*, & Matthews, T.J. (1991). Alteration of HIV-1 infectivity and neutralization by a single amino acid replacement in the V3 loop domain. *AIDS Res. Hum. Retroviruses* **7**, 595-603.
15. Hwang, S.S., Boyle, T.J., Lyerly, H.K. & Cullen, B.R. (1992). Identification of envelope V3 loop as the major determinant of CD4 neutralization sensitivity of HIV-1. *Science* **257**, 535-537.
16. Bhattacharyya, D., Brooks, B.R. & Callahan, L. (1996). Positioning of positively charged residues in the V3 loop correlates with HIV type 1 syncytium-inducing phenotype. *AIDS Res. Hum. Retroviruses* **12**, 83-90.
17. Speck, R.F., *et al.*, & Goldsmith, M.A. (1997). Selective employment of chemokine receptors as human immunodeficiency virus type 1 coreceptors determined by individual amino acids within the envelope V3 loop. *J. Virol.* **71**, 7136-7139.
18. De Jong, J.J., De Ronde, A., Keulen, W., Tersmette, M. & Goudsmit, J. (1992). Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. *J. Virol.* **66**, 6777-6780.
19. Fouchier, R.A., *et al.*, & Schuitemaker, H. (1992). Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J. Virol.* **66**, 3183-3187.
20. Wyatt, R., *et al.*, & Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**, 705-711.
21. Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. & Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648-659.
22. Rizzuto, C.D., *et al.*, & Sodroski, J. (1998). A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science* **280**, 1949-1953.
23. Wilmot, C.M. & Thornton, J.M. (1988). Analysis and prediction of the different types of β-turn in proteins. *J. Mol. Biol.* **203**, 221-232.
24. White-Scharf, M.E., Potts, B.J., Smith, L.M., Sokolowski, K.A., Rusche, J.R. & Silver, S. (1993). Broadly neutralizing monoclonal antibodies to the V3 region of HIV-1 can be elicited by peptide immunization. *Virology* **192**, 197-206.
25. Valenzuela, A., Blanco, J., Krust, B., Franco, R. & Hovanessian, A.G. (1997). Neutralizing antibodies against the V3 loop of human immunodeficiency virus type 1 gp120 block the CD4-dependent and -independent binding of virus to cells. *J. Virol.* **71**, 8289-8298.
26. Rini, J.M., Stanfield, R.L., Stura, E.A., Salinas, P.A., Profy, A.T. & Wilson, I.A. (1993). Crystal structure of a human immunodeficiency virus type 1 neutralizing antibody, 50.1, in complex with its V3 loop peptide antigen. *Proc. Natl Acad. Sci. USA* **90**, 6325-6329.
27. Stanfield, R.L., Takimoto-Kamimura, M., Rini, J.M., Profy, A.T. & Wilson, I.A. (1993). Major antigen-induced domain rearrangements in an antibody. *Structure* **1**, 83-93.
28. Ghiara, J.B., Stura, E.A., Stanfield, R.L., Profy, A.T. & Wilson, I.A. (1994). Crystal structure of the principal neutralization site of HIV-1. *Science* **264**, 82-85.
29. Ghiara, J.B., Ferguson, D.C., Satterthwait, A.C., Dyson, H.J. & Wilson, I.A. (1997). Structure-based design of a constrained peptide mimic of the HIV-1 V3 loop neutralization site. *J. Mol. Biol.* **266**, 31-39.
30. Burgess, A.W. & Leach, S.J. (1973). An obligatory α-helical amino acid residue. *Biopolymers* **12**, 2599-2605.
31. Marshall, G.R. & Bosshard, H.E. (1972). Angiotensin II. Studies on the biologically active conformation. *Circ. Res.* **31**, Supplement 2, 143-150.
32. Brünger, A.T., Krukowski, A. & Erickson, J.W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. A* **46**, 585-593.
33. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47**, 392-400.
34. Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest.* US Department of Health and Human Services.
35. Ratner, L., *et al.*, & Wong-Staal, F. (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**, 277-284.
36. Laskowski, R.A., MacArthur, M.W., Hutchinson, E.G. & Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364.
37. Stanfield, R.L., Fieser, T.M., Lerner, R.A. & Wilson, I.A. (1990). Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 Å. *Science* **248**, 712-719.
38. Arevalo, J.H., Stura, E.A., Taussig, M.J. & Wilson, I.A. (1993). Three-dimensional structure of an anti-steroid Fab' and progesterone–Fab' complex. *J. Mol. Biol.* **231**, 103-118.
39. Al-Lazikani, B., Lesk, A.M. & Chothia, C. (1997). Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.* **273**, 927-948.
40. Martin, A.C. & Thornton, J.M. (1996). Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. *J. Mol. Biol.* **263**, 800-815.
41. Brünger, A.T., Leahy, D.J., Hynes, T.R. & Fox, R.O. (1991). 2.9 Å Resolution structure of an anti-dinitrophenyl-spin-label monoclonal antibody Fab fragment with bound hapten. *J. Mol. Biol.* **221**, 239-256.
42. Bossart-Whitaker, P., Chang, C.Y., Novotny, J., Benjamin, D.C. & Sheriff, S. (1995). The crystal structure of the antibody N_{10}-staphylococcal nuclease complex at 2.9 Å resolution. *J. Mol. Biol.* **253**, 559-575.
43. Marquart, M., Deisenhofer, J., Huber, R. & Palm, W. (1980). Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen-binding fragment at 3.0 Å and 1.9 Å resolution. *J. Mol. Biol.* **141**, 369-391.
44. He, X.M., Ruker, F., Casale, E. & Carter, D.C. (1992). Structure of a human monoclonal antibody Fab fragment against gp41 of human immunodeficiency virus type 1. *Proc. Natl Acad. Sci. USA* **89**, 7154-7158.
45. Altschuh, D., Vix, O., Rees, B. & Thierry, J.C. (1992). A conformation of cyclosporin A in aqueous environment revealed by the X-ray structure of a cyclosporin–Fab complex. *Science* **256**, 92-94.
46. Shirai, H., Kidera, A. & Nakamura, H. (1996). Structural classification of CDR-H3 in antibodies. *FEBS Lett.* **399**, 1-8.
47. Connolly, M.L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science* **221**, 709-713.
48. Wilson, I.A., *et al.*, & Stura, E.A. (1991). Structural aspects of antibodies and antibody–antigen complexes. *Ciba Foundation Symposium.* **159**, 13-28; discussion 28-39.
49. Stanfield, R.L. & Wilson, I.A. (1993). X-ray crystallographic studies of antibody–peptide complexes. *Immunomethods* **3**, 211-221.
50. Jellis, C.L., *et al.*, & Gray, G.S. (1993). Defining critical residues in the epitope for a HIV-neutralizing monoclonal antibody using phage display and peptide array technologies. *Gene* **137**, 63-68.
51. Catasti, P., Fontenot, J.D., Bradbury, E.M. & Gupta, G. (1995). Local and global structural properties of the HIV-MN V3 loop. *J. Biol. Chem.* **270**, 2224-2232.

52. Gupta, G., Anantharamaiah, G.M., Scott, D.R., Eldridge, J.H. & Myers, G. (1993). Solution structure of the V3 loop of a Thailand HIV isolate. *J. Biomol. Struct. Dyn.* **11**, 345-366.
53. Chandrasekhar, K., Profy, A.T. & Dyson, H.J. (1991). Solution conformational preferences of immunogenic peptides derived from the principal neutralizing determinant of the HIV-1 envelope glycoprotein gp120. *Biochemistry* **30**, 9187-9194.
54. Vu, H.M., de Lorimier, R., Moody, M.A., Haynes, B.F. & Spicer, L.D. (1996). Conformational preferences of a chimeric peptide HIV-1 immunogen from the C4-V3 domains of gp120 envelope protein of HIV-1 CAN0A based on solution NMR: comparison to a related immunogenic peptide from HIV-1 RF. *Biochemistry* **35**, 5158-5165.
55. Vranken, W.F., *et al.*, & Borremans, F.A. (1996). Conformational features of a synthetic cyclic peptide corresponding to the complete V3 loop of the RF HIV-1 strain in water and water/trifluoroethanol solutions. *Eur. J. Biochem.* **236**, 100-108.
56. Sarma, A.V., Raju, T.V. & Kunwar, A.C. (1997). NMR study of the peptide present in the principal neutralizing determinant (PND) of HIV-1 envelope glycoprotein gp120. *J. Biochem. Biophys. Methods* **34**, 83-98.
57. Tolman, R.L., *et al.*, & Conley, A.J. (1993). Cyclic V3-loop-related HIV-1 conjugate vaccines. Synthesis, conformation and immunological properties. *Int. J. Pept. Protein Res.* **41**, 455-466.
58. Jelinek, R., Terry, T.D., Gesell, J.J., Malik, P., Perham, R.N. & Opella, S.J. (1997). NMR structure of the principal neutralizing determinant of HIV-1 displayed in filamentous bacteriophage coat protein. *J. Mol. Biol.* **266**, 649-655.
59. Huang, X., Smith, M.C., Berzofsky, J.A. & Barchi, J.J., Jr. (1996). Structural comparison of a 15 residue peptide from the V3 loop of HIV-1IIIb and an O-glycosylated analogue. *FEBS Lett.* **393**, 280-286.
60. Huang, X., *et al.*, & Garrity, R.R. (1997). Glycosylation affects both the three-dimensional structure and antibody binding properties of the HIV-1IIIB GP120 peptide RP135. *Biochemistry* **36**, 10846-10856.
61. Markert, R.L., *et al.*, & Griesinger, C. (1996). Secondary structural elements as a basis for antibody recognition in the immunodominant region of human immunodeficiency viruses 1 and 2. *Eur. J. Biochem.* **237**, 188-204.
62. Zvi, A., Kustanovich, I., Hayek, Y., Matsushita, S. & Anglister, J. (1995). The principal neutralizing determinant of HIV-1 located in V3 of gp120 forms a 12-residue loop by internal hydrophobic interactions. *FEBS Lett.* **368**, 267-270.
63. Zvi, A., Feigelson, D.J., Hayek, Y. & Anglister, J. (1997). Conformation of the principal neutralizing determinant of human immunodeficiency virus type 1 in complex with an anti-gp120 virus neutralizing antibody studied by two-dimensional nuclear magnetic resonance difference spectroscopy. *Biochemistry* **36**, 8619-8627.
64. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.
65. Housset, D., Habersetzer-Rochat, C., Astier, J.P. & Fontecilla-Camps, J.C. (1994). Crystal structure of toxin II from the scorpion *Androctonus australis* Hector refined at 1.3 Å resolution. *J. Mol. Biol.* **238**, 88-103.
66. Mande, S.C., Mehra, V., Bloom, B.R. & Hol, W.G. (1996). Structure of the heat shock protein chaperonin-10 of *Mycobacterium leprae*. *Science* **271**, 203-207.
67. Achour, A., *et al.*, & Karre, K. (1998). The crystal structure of H-2D^d MHC class I complexed with the HIV-1-derived peptide P18-I10 at 2.4 Å resolution: implications for T cell and NK cell recognition. *Immunity* **9**, 199-208.
68. Brünger, A.T. (1990). Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. *Acta Crystallogr. A* **46**, 46-57.
69. Harada, Y., Lifchitz, A., Berthou, J. & Jolles, P. (1981). A translation function combining packing and diffraction information: an application to lysozyme (high-temperature form). *Acta Crystallogr. A* **37**, 398-406.
70. Jones, T.A. (1985). Diffraction methods for biological macromolecules. Interactive computer graphics: FRODO. *Methods Enzymol.* **115**, 157-171.
71. Israel, M., Chirino, A.J. & Schuller, D.J. (1996). TOM/FRODO, Alberta/Caltech Version 3.3.
72. Brünger, A.T. (1992). *X-PLOR Version 3.1*. Yale University Press, New Haven, CT, USA.
73. Fitzgerald, P.M.D. (1988). MERLOT, an integrated package of computer programs for the determination of crystal structures by molecular replacement. *J. Appl. Crystallogr.* **21**, 273-278.
74. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. *J. Appl. Crystallogr.* **24**, 946-950.
75. Esnouf, R.M. (1997). An extensively modified version of MOLSCRIPT that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 112-113, 132-134.
76. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D Version 2.0 – a program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869-873.
77. Bacon, D.J. & Anderson, W.F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219-220.
78. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296.
79. Ferrin, T.E., Huang, C.C., Jarvis, L.E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graph.* **6**, 13-27.
80. Huang, C.C., Pettersen, E.F., Klein, T.E., Ferrin, T.E. & Langridge, R. (1991). Conic: a fast renderer for space-filling molecules with shadows. *J. Mol. Graph.* **9**, 230-236.
81. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
82. Xuong, N.-H., Nielsen, C., Hamlin, R. & Anderson, D. (1985). Strategy for data collection from protein crystals using a multiwire counter area detector diffractometer. *J. Appl. Crystallogr.* **18**, 342-351.
83. Howard, A.J., Nielsen, C. & Xuong, N.H. (1985). Software for a diffractometer with multiwire area detector. *Methods Enzymol.* **114**, 452-472.
84. Sheriff, S., Hendrickson, W.A. & Smith, J.L. (1987). Structure of myohemerythrin in the azidomet state at 1.7/1.3 Å resolution. *J. Mol. Biol.* **197**, 273-296.
85. Gelin, B.R. & Karplus, M. (1979). Side-chain torsional potentials: effect of dipeptide, protein, and solvent environment. *Biochemistry* **18**, 1256-1268.
86. McDonald, I., & Thornton, J. (1995) The application of hydrogen bonding analysis in X-ray crystallography to help orientate asparagine, glutamine and histidine side chains. *Protein Eng.* **8**, 217-224.

Because *Structure with Folding & Design* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed (accessed from http://biomednet.com/cbiology/str). For further information, see the explanation on the contents page.

The Hydrogen Bond Mimic Approach: Solid-Phase Synthesis of a Peptide Stabilized as an α-Helix with a Hydrazone Link

Edelmira Cabezas[†] and Arnold C. Satterthwait*,[†]

Contribution from the Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037

Received September 8, 1998. Revised Manuscript Received February 15, 1999

Abstract: Proteins are characterized by extensive hydrogen bonding that defines regular and irregular substructures. However, hydrogen bonds are weak and insufficient for stabilizing peptide conformation in water. Consequently, the biological activity of peptides is reduced. This led us to test whether a covalent mimic of the hydrogen bond could be used to stabilize peptide conformation in water. A solid-phase synthesis is described for replacing a main-chain hydrogen bond (NH → O=CRNH) with a hydrazone link (N−N=CH−CH_2CH_2) in peptides. The synthesis is easy to implement, rapid, and capable of high yields. The replacement of a putative (i + 4 → i) hydrogen bond with the hydrazone at the N terminus of acetyl-GLAGAEAAKA-NH_2 (**1**) to give [JLAZ]AEAAKA-NH_2 (**2**) converts it to a full-length α-helix in water at ambient temperature as indicated by NMR spectroscopy. The observation of weak $d_{\alpha N}(i, i + 3)$, medium $d_{NN}(i, i + 1)$, and strong $d_{\alpha\beta}(i, i + 3)$ NOEs that span **2** establish the formation of a full-length α-helix in water. $J_{\alpha N}$ coupling constants and amide proton chemical shifts and temperature coefficients are consistent with a model involving rapidly equilibrating extended and α-helical conformers. Substituting L-alanine with L-proline to give [JLPZ]AEAAKA-NH_2 (**3**) enhances α-helix nucleation and shifts the equilibrium further toward full-length α-helix. The hydrazone link displays many of the properties required of a hydrogen bond mimic and could find use as a general means for constraining peptides to a range of biologically relevant conformations.

Introduction

Synthetic peptides are a rich source of drug[1,2] and vaccine leads.[3] Biologically active peptides are often searched for by screening peptides derived from proteins[4] or by screening peptide libraries[5] against receptors. However, most peptide fragments from proteins are conformationally heterogeneous in aqueous solution.[6] Since receptors bind peptides in specific conformations, it is important to consider the effect that peptide conformation has on binding energy. This question, first treated formally by Anfinsen,[7] led to the conclusion that the binding energy of a peptide is directly proportional to the free energy required for folding the peptide into the receptor bound conformation. The free-energy barrier to folding a peptide can be considerable. When synthetic peptides are constrained to shapes that mirror the binding pockets of receptors, the energy barrier is reduced, and 10^2−10^5-fold improvements in affinities are achieved.[7−9] Differences of this magnitude are sufficient for identifying new peptide leads and receptors[10] and for

* Address correspondence to this author at the following address: The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, California 92037. Telephone: 619-646-3100, ext. 3658. Fax: 619-646-3195. E-mail: asat@burnham-inst.org.

† Present address: The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037.

(1) Hruby, V. J. *Life Sci.* **1982**, *31*, 189−199. Hruby, V. J.; Al-Obeidi, F.; Kazmierski, W. *Biochem. J.* **1990**, *268*, 249−262. Li, B.; Tom, J. Y. K.; Oare, D.; Yen, R.; Fairbrother, W. J.; Wells, J. A.; Cunningham, B. C. *Science* **1995**, *270*, 1657−1660. Wrighton, N. C.; Farrell, F. X.; Chang, R.; Kashyap, A. K.; Barbone, F. P.; Mulcahy, L. S.; Johnson, D. L.; Barrett, R. W.; Jolliffe, L. K.; Dower, W. J. *Science* **1996**, *273*, 458−464.

(2) Müller, K.; Obrecht, D.; Knierzinger, A.; Stankovic, C.; Spiegler, C.; Bannwarth, W.; Trzeciak, A.; Englert, G.; Labhardt, A. M.; Schönholzer, P. *Perspect. Med. Chem.* **1993**, 513−531.

(3) (a) Sela, M.; Schechter, B.; Schechter, I.; Borek, F. *Cold Spring Harbor Symp. Quant. Biol.* **1967**, *39*, 537−545. (b) Lerner, R. A. *Adv. Immunol.* **1984**, *36*, 1−44. (c) Brown, F. *Philos. Trans. R. Soc. London Ser. B* **1994**, *344*, 213−219.

(4) Geysen, H. M.; Meloen R. H.; Barteling, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3998−4002. Houghten, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5131−5135.

(5) (a) Scott, J. K.; Smith, G. P. *Science* **1990**, *249*, 386−390. (b) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−84. (c) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84−86.

(6) (a) Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 795−817. (b) Dyson, H. J.; Sayre, J. R.; Merutka, G.; Shin, H. C.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 819−835.

(7) Sachs, D. H.; Schechter, A. N.; Eastlake, A.; Anfinsen C. B. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 3790−3794. Anfinsen, C. B. *Science* **1973**, *181*, 223−229.

(8) (a) Teicher, E.; Maron, E.; Arnon, R. *Immunochemistry* **1973**, *10*, 265−271. (b) Furie, B.; Schechter, A. N.; Sachs, D. H.; Anfinsen, C. B. *J. Mol. Biol.* **1975**, *92*, 497−506. (c) Komoriya, A.; Hortsch, M.; Meyers, C.; Smith, M.; Kanety, H.; Schlessinger, J. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1351−1355. (d) Judice, J. K.; Tom, J. Y. K.; Huang, W.; Wrin, T.; Vennari, J.; Petropoulos, C. X. J.; McDowell, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13426−13430.

(9) (a) Satterthwait, A. C.; Cabezas, E.; Calvo, J. C.; Wu, J. X.; Wang, P. L.; Chen, S. Q.; Kaslow, D. C.; Livnah, O.; Stura, E. A. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fourteenth American Peptide Symposium;* Mayflower Scientific Ltd.: England, 1996; pp 772−773. (b) Cabezas, E.; Stanfield, R.; Wilson, I. A.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fourteenth American Peptide Symposium;* Mayflower Scientific Ltd.: England, 1996; pp 800−801.

(10) (a) Calvo, J. C.; Perkins, M.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Thirteenth American Peptide Symposium;* ESCOM: Leiden, 1994; pp 725−726. (b) Satterthwait, A. C.; Cabezas, E.; Calvo, J. C.; Chen, S. Q.; Wu, J. X.; Wang, P. L.; Xie, Y. L.; Stura, E. A.; Kaslow, D. C. *Peptides: Biology and Chemistry; Proceedings of the 1994 Chinese Peptide Symposium;* ESCOM: Leiden, 1995; pp 229−233.

10.1021/ja983212t CCC: $18.00
Published on Web 04/07/1999

converting immunologically inactive peptides into active peptides.[10,11] Conversely, this implies that many peptide leads and receptors are lost because peptides unfold. Synthetic vaccine development may be even more dependent on the mimicry of protein structure.[3a,9-12] Consequently, there are compelling reasons for developing methods for constraining peptides to the substructures occupied by their cognate sequences in proteins.

To address this need, we proposed the development of covalent hydrogen bond mimics.[13] On average, more than 60% of the amino acids in globular proteins form intramolecular hydrogen bonds between main-chain peptide bonds.[14] In addition, secondary and irregular structures can be defined by different hydrogen-bonding patterns.[15] We therefore reasoned that the replacement of putative structure-defining hydrogen bonds with a covalent mimic could yield a general approach for constraining peptides to different protein substructures based on the position of the link and the amino acid sequence.

Initially, Arrhenius and Satterthwait[16] developed a solution-phase synthesis for replacing a main-chain hydrogen bond with a hydrazone link. However, the synthesis was lengthy, yields were low, and solubility problems arose with fully protected peptides. In this paper, we report a synthetic protocol for the insertion of the hydrazone link into a peptide by solid-phase synthesis. It allowed us to explore the α-helix-stabilizing potential of the hydrazone link in water. Work on medium-sized loops is being reported separately.[9,17]

The α-helix is an important structure in proteins.[15a,d,18] It presents a definitive test for any main-chain hydrogen bond mimic since structural incompatibilities are more likely to become evident in the α-helix than in less crowded structures such as loops. Several synthetic approaches to helix stabilization have been taken.[19] One approach builds on templates to propagate helices.[2,20,21] A second approach links amino acid side chains in a position-dependent manner to stabilize helices.[22] A third approach uses linkers, bridges or synthetic scaffolds to position peptide chains in proximity for helix-stabilizing interactions.[23]

Table 1. Representative Regular and Irregular Protein Structures Defined by Main-Chain to Main-Chain Amide Hydrogen Bonds[a,b]

protein substructure	hydrogen bond	ring size
3_{10}-helix; β-turns	$i+3 \rightarrow i$	10
β-hairpin loops, $n > 0$	$i+3+n \rightarrow i$	$10+3n$
α-helix	$i+4 \rightarrow i$	13
3:3 loop (rare:irregular)		
2:2 loop (defined)	$i \rightarrow i+3$	14
2:4 loop (rare, irregular)		
β-hairpin loops, $n > 0$	$i \rightarrow i+3+n$	$14+3n$
4:4 loop (defined and irregular)	$i+5 \rightarrow i$	16
3:5 loop (defined and irregular)	$i \rightarrow i+4$	17
4:6 loop (irregular)	$i \rightarrow i+5$	20
antiparallel β-sheet, parallel β-sheet	$i \rightarrow i+n$ $i+n \rightarrow i$	variable or discontinuous

[a] Reference 15. [b] The hydrogen bond is indicated by an arrow. The arrow points from the residue donating the amide NH proton to the residue bearing the accepting carbonyl oxygen atom (NH → O=CRNH). The reference residue, *i*, is N-terminal to the second residue which is separated by the indicated number of amino acids, *n*. The ring size is defined by the number of atoms connected by the hydrogen bond.

Design

A review of hydrogen bonding in proteins[14,15] shows that different substructures can be defined by their hydrogen-bonding patterns between main-chain peptide bonds on the basis of ring size and orientation (Table 1). The hydrazone link (N−N=CH−CH_2CH_2) was designed to replace the hydrogen bond, (NH → O=C(R)NH) that forms between the main-chain amide proton of an upstream amino acid and the main-chain carbonyl oxygen of a downstream amino acid ($i + n \rightarrow i$). The hydrogen bond is replaced by the N=CH double bond, and the associated

(11) Lairmore, M. D.; Lal, R. B.; Kaumaya, P. T. *Biomed. Pept. Proteins Nucleic Acids* **1995**, *1*, 117−122. Valero, M. L.; Camarero, J. A.; Adeva, A.; Verdaguer, N.; Fita, I.; Mateu, M. G.; Domingo, E.; Giralt, E.; Andreu, D. *Biomed. Pept. Proteins Nucleic Acids* **1995**, *1*, 133−140. Sheth, H. D.; Glasier, L. M.; Ellert, N. W.; Cachia, P.; Kohn, W.; Lee, K. K.; Paranchych, W.; Hodges, R. S.; Irvin, R. T. *Biomed. Pept. Proteins Nucleic Acids* **1995**, *1*, 141−148. Urban, J.; Qabar, M.; Sia, C.; Klein, M.; Kahn, M. *Bioorg. Med. Chem.* **1996**, *4*, 673−676

(12) (a) Jemmerson, R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 9180−9184. (b) Satterthwait, A. C.; Arrhenius, T.; Hagopian, R. A.; Zavala, F.; Nussenzweig, V.; Lerner, R. A. *Vaccine* **1988**, *6*, 99−103. (c) Laver, W. G.; Air, G. M.; Webster, R. G.; Smith-Gill, S. J. *Cell* **1990**, *61*, 553−556. (d) Stanfield, R. L.; Feiser, T. M.; Lerner, R. A.; Wilson, I. A. *Science* **1990**, *248*, 712−719. (e) Müller, S.; Plaue, S.; Samama, J. P.; Valette, M.; Briand, J. P.; Van Regenmortel, M. H. V. *Vaccine* **1990**, *8*, 308−314. (f) Spangler, B. D. *J. Immunol.* **1991**, *146*, 1591−1595. (g) Cachia, P. J.; Glasier, L. M.; Hodgins, R. R.; Wong, W. Y.; Irvin, R. T.; Hodges, R. S. *J. Pept. Res.* **1998**, *52*, 289−299. (h) Verdaguer, N.; Sevilla, N.; Valero, M. L.; Stuart, D.; Brocchi, E.; Andreu, D.; Giralt, E.; Domingo, E.; Mateu, M. G.; Fita, I. *J. Virol.* **1998**, *72*, 739−748

(13) Arrhenius, T.; Lerner, R. A.; Satterthwait, A. C. *Protein Structure, Folding and Design 2;* Alan R. Liss, Inc.: New York, 1987; pp 453−465.

(14) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.

(15) (a) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167−339. (b) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221−232. Wilmot, C. M.; Thornton, J. M. *Protein Eng.* **1990**, *3*, 479−493. (c) Sibanda, B. L.; Blundell, T. L.; Thornton, J. M. *J. Mol. Biol.* **1989**, *206*, 759−777. (d) Richardson, J. S.; Richardson, D. C. *Prediction of Protein Structure and the Principles of Protein Conformation;* Plenum Press: New York, 1989; pp 1−98. (e) Donate, L. E.; Rufino, S. D.; Canard, L. H. J.; Blundell, T. L. *Protein Sci.* **1996**, *5*, 2600−2616.

(16) Arrhenius, T.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Eleventh American Peptide Symposium;* ESCOM: Leiden, 1990; pp 870−872.

(17) (a) Cabezas, E.; Wang, P. L.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fourteenth American Peptide Symposium;* Mayflower Scientific Ltd.: England, 1996; pp 734−735. (b) Cabezas, E.; Satterthwait, A. C. *Peptides: Chemistry, Structure and Biology; Proceedings of the Fifteenth American Peptide Symposium;* ESCOM: Leiden, in press. (c) Stanfield, R. L.; Cabezas, E.; Satterthwait, A. C.; Stura, E. A.; Profy, A. T.; Wilson, I. A. *Structure*, **1999**, *7*, 131−142.

(18) Kaiser, E. T.; Kezdy, F. J. *Science* **1984**, *223*, 249−255. Degrado, W. F. *Adv. Protein Chem.* **1988**, *39*, 51−124.

(19) *Houben-Weyl Methods in Organic Chemistry. Synthesis of Peptides and Peptidomimetics;* Georg Thieme Verlag: Stuttgart, 1999; Vol. E22b, manuscript in preparation.

(20) (a) Kemp, D. S.; Curran, T. P. *Tetrahedron Lett.* **1988**, *29*, 4935−4938. (b) Kemp, D. S.; Curran, T. P.; Davis, W. M.; Boyd, J. G.; Muendel, C. *J. Org. Chem.* **1991**, *56*, 6672−6682. (c) Kemp, D. S.; Curran, T. P.; Boyd, J. G.; Allen, T. J. *J. Org. Chem.* **1991**, *56*, 6683−6697. (d) Kemp, D. S.; Allen, T. J.; Oslick, S. L. *J. Am. Chem. Soc.* **1995**, *117*, 6641−6657. (e) Kemp, D. S.; Allen, T. J.; Oslick, S. L.; Boyd, J. G. *J. Am. Chem. Soc.* **1996**, *118*, 4240−4248. (f) Kemp, D. S.; Rothman, J. H. *Tetrahedron Lett.* **1995**, *36*, 4023−4026.

(21) Austin, R. E.; Maplestone, R. A.; Sefler, A. M.; Liu, K.; Hruzewicz, W. N.; Liu, C. W.; Cho, H. S.; Wemmer, D. E.; Bartlett, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 6461−6472.

(22) (a) Felix, A. M.; Heimer, E. P.; Wang, C. T.; Lambros, T. J.; Fournier, A.; Mowles, S. M.; Maines, S.; Campbell, R. M.; Wegrzynski, B. B.; Toome, V.; Fry, D.; Madison, V. S. *Int. J. Pept. Protein Res.* **1988**, *32*, 441−454. (b) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1990**, *112*, 6046−6051. (c) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1990**, *114*, 6966−6973. (d) Bracken, C.; Gulyás, J.; Taylor, J. W.; Baum, J. *J. Am. Chem. Soc.* **1994**, *116*, 6431−6432. (e) Houston, M. E.; Gannon, C. L.; Kay, C. M.; Hodges, R. S. *J. Pept. Sci.* **1995**, *1*, 274−282. (f) Jackson, D. Y.; King, D. S.; Chmielewski, J.; Singh, S.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 9391−9392. (g) Yu, C.; Taylor, J. W. *Tetrahedron Lett.* **1996**, *37*, 1731−1734. (h) Phelan, J. C.; Skelton, N. J.; Braisted, A. C.; McDowell, R. S. *J. Am. Chem. Soc.* **1997**, *119*, 455−460.

(23) Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535−554. Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825−831. Dawson, P. E.; Kent, S. B. H. *J. Am. Chem. Soc.* **1993**, *115*, 7263−7266. Monera, O. D.; Sonnichsen, F. D.; Hicks, L.; Kay, C. M.; Hodges, R. S. *Protein Eng.* **1996**, *9*, 353−363. Mant, C. T.; Chao, H.; Hodges, R. S. *J. Chromatog. A* **1997**, *791*, 85−98. Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. *Protein Sci.* **1998**, *7*, 1404−1414. Ghirlanda, G.; Lear, J. D.; Lombardi, A.; DeGrado, W. F. *J. Mol. Biol.* **1998**, *281*, 379−391.

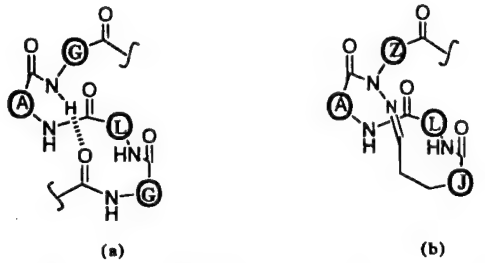


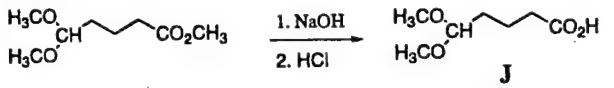
Figure 1. The substitution of an ($i + 4 \rightarrow i$) hydrogen bond in one turn of an α-helix (a) with a hydrazone covalent hydrogen bond mimic (b). The depicted structure (b) is [JLAZ]-. Circled letters at the α-carbons are one letter codes for amino acids and modified amino acids used at these positions. L replaces $HC(CH_2)CH(CH_3)_2$. A replaces $HC(CH_3)$. G, J, and Z replace CH_2.

peptide link is replaced with an ethylene group (Figure 1). Although neither hydrogen bond length nor angles are precisely mimicked by the double bond,[24] the attached hydrocarbon chain provides enough flexibility for the hydrazone link to substitute for structure defining hydrogen bonds in molecular models of peptides in many different conformations.

The α-helix is defined by a pattern of sequential ($i + 4 \rightarrow i$) hydrogen bonds.[14,15a,d] When an N-terminal ($i + 4 \rightarrow i$) hydrogen bond is replaced with the hydrazone link, a 13-membered ring is formed that corresponds to the 13-membered hydrogen-bonded turn of an α-helix (Figure 1). The hydrazone is formed by the reaction of an activated acetal (**J**) with a hydrazino derivative (**Z**). The hydrazone link fits compactly into a space-filling Corey−Pauling−Koltum molecular model of an α-helical peptide. **Z** is restricted to an α-methylene carbon (glycine equivalent) to alleviate a steric clash of side chains with the −N=CH− bond that would occur in an α-helix.

The hydrazone link was substituted for the putative ($i + 4 \rightarrow i$) hydrogen bond at the N terminus of a linear peptide, acetyl-GLAGAEAAKA-NH_2 (**1**), to give [JLAZ]AEAAKA-NH_2 (**2**). [JLAZ]- is depicted in Figure 1. The constrained peptide should form about three turns of an α-helix which is about the average size for an α-helix in a globular protein.[25a] The alanine-rich extension, -AEAAKA-NH_2 was selected on the basis of its potential for α-helix formation and disinclination to aggregate.[26] The conformations of these two peptides were compared using principally NMR spectroscopy. Alanine-based peptides are preferred for NMR studies with helical peptides because alanine is both a helix-stabilizing amino acid[26] and the only amino acid with equivalent β protons. Thus, distances to these protons are independent of side-chain rotamer distributions and consequently predictable for specific conformations. Since α-helices are characterized by a unique pattern of short, sequential distances $d_{\alpha\beta}(i, i + 3)$, alanine-based peptides can be used to test for the α-helical conformation on a per residue basis.[20e] Although NMR signals for alanine often overlap, they can be edited and assigned by means of deuterated alanine analogues.[20e] A second constrained peptide, [JLPZ]AEAAKA-NH_2 (**3**), was synthesized to test whether L-proline with $\phi = -60°$,[15d] close to the ideal for an α-helix ($\phi = -57°$), could enhance α-helix stabilization. L-proline often initiates α-helix formation in proteins.[25]

Scheme 1



Scheme 2

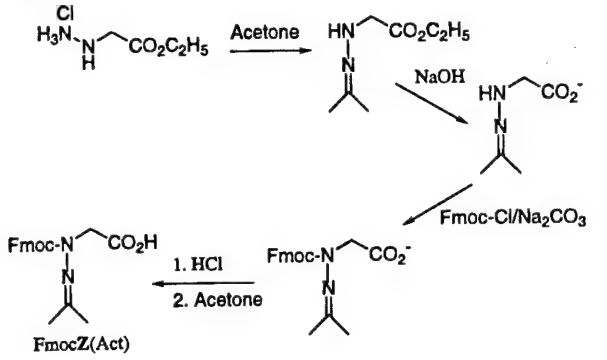


Table 2. Linear and Constrained Peptides Prepared for This Study[a]

1:	acetyl-G1-L2-A3-G4-A5-E6-A7-A8-K9-A10-NH_2
1a:	acetyl-GLA^{1}GAEAAKA-NH_2
1b:	acetyl-GLAGAEA^{1}AKA-NH_2
1c:	acetyl-GLAGAEAA^{1}KA-NH_2
1d:	acetyl-GLAGAEAAKA^1-NH_2
2:	[J1-cL2-cA3-Z4]-A5-E6-A7-A8-K9-A10-NH_2
2a:	[JLAZ]A^{4}EA^{4}A^{1}KA^4-NH_2
2b:	[JLAZ]A^{3}EA^{4}A^{1}KA^4-NH_2
2c:	[JLAZ]A^{4}EA^{3}A^{4}KA^1-NH_2
3:	[J1-cL2-cP3-Z4]-A5-E6-A7-A8-K9-A10-NH_2

[a] Amino acids (in bold) are in one-letter code and numbered according to their position in the sequence: c indicates the amino acid is from the cyclized segment of the peptide. **J** and **Z** are identified in Schemes 1, 2. [JLAZ] is the cyclized peptide in Figure 1b. A^1 is L-alanine-2-d_1; A^3 is L-alanine-3,3,3-d_3; A^4 is L-alanine-2,3,3,3-d_4.

Synthesis

The insertion of the hydrazone link into peptides by solid-phase peptide synthesis requires the prior preparation of 5,5-dimethoxy-1-oxopentanoic acid (**J**), (1-methyethylidene-2-Fmoc)hydrazinoacetic acid (Fmoc-**Z**(Act) and an *N*-α-Fmoc-L-amino acid chloride. **J** was prepared as the methyl ester in three steps according to Stevens and Lee[27] and converted to the acid (Scheme 1). The acid is stable as a neat liquid at −20 °C. Fmoc-**Z**(Act) was prepared in a one-pot synthesis according to Scheme 2. *N*-α-Fmoc-L-amino acid chlorides were prepared from the *N*-α-Fmoc-amino acid and thionyl chloride according to Carpino et al.[28]

A list of peptides synthesized for this study is in Table 2. Linear peptides including **1** were synthesized using standard Fmoc synthesis. [JLAZ]AEAAKA-NH_2 (**2**) and its deuterated analogues were synthesized on Rink's amide resin according to Scheme 3. Standard Fmoc chemistry was used for coupling *N*-α-Fmoc-amino acids, Fmoc-**Z**(Act) and **J**. The most critical reaction in peptide assembly was the coupling of *N*-α-Fmoc-L-alanine to the secondary amine, **Z**(Act). Several coupling procedures were tested including prior reaction of *N*-α-Fmoc-L-alanine with EDC to form the anhydride and activation of

(24) The mean hydrogen bond length and angles for parallel and antiparallel β-sheets and α-helix are NH···O, 2.0 Å; NH···O=C, 155°; N−H···O, 160°; Chothia, C. *Annu. Rev. Biochem.* **1984**, *53*, 537−572. The C=N bond length for pyridine, etc. is 1.34 Å; Gordon, A. J.; Ford, R. A. *The Chemists Companion; A Handbook of Practical Data, Techniques, and References;* John Wiley: New York, 1972; p 108.

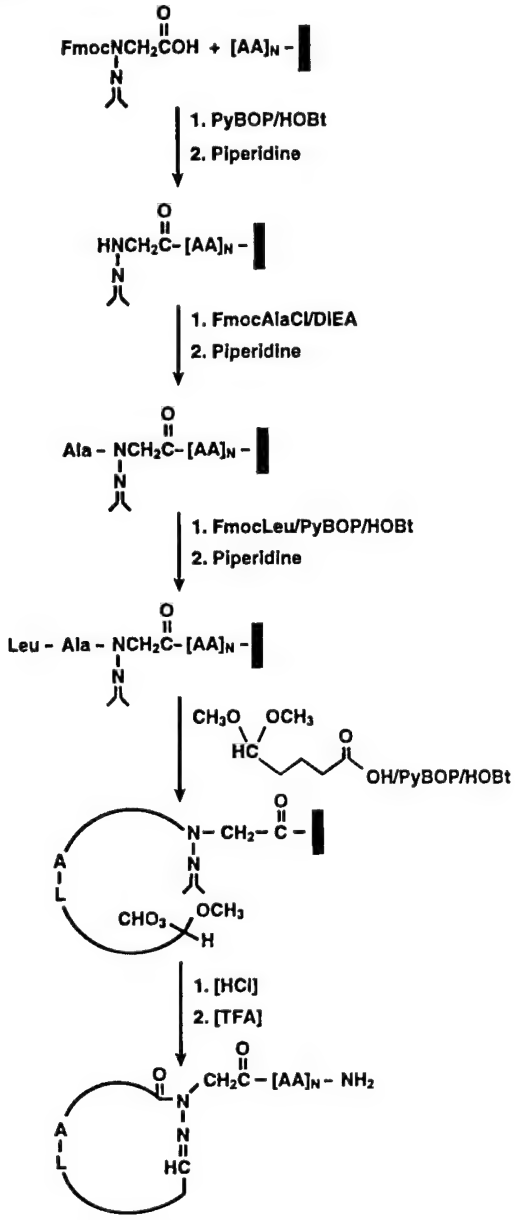
(25) (a) Presta, L. G.; Rose, G. D.*Science* **1988**, *240*, 1632−1641. (b) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648−1652.

(26) (a) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8898−8902. (b) Scholtz, J. M.; Marqusee, S.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Santoro, M.; Bolen, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2854−2858.

(27) Stevens, R. V.; Lee, A. W. M. *J. Am. Chem. Soc.* **1979**, *101*, 7032−7035.

(28) Carpino, L. A.; Cohen, B. J.; Stephens, K. E., Jr.; Sadat-Aalaee, S. Y.; Tien, J. H.; Langridge, D. C. *J. Org. Chem.* **1986**, *51*, 3732−3734.

Scheme 3



N-α-Fmoc-L-alanine with PyBOP/HOBt or DIC/HOBt. However, none of these methods proved as effective as coupling with *N*-α-Fmoc-alanine chloride. For the most efficient coupling, the acyl chloride in *N*-methylpyrrolidone was preequilibrated with **Z**(Act)-Rink-resin before adding diisopropylethylamine. The coupling reaction is complete within 15 min. Either premixing the acyl chloride with diisopropylethylamine or adding the base to **Z**(Act)-Rink-resin prior to the acyl chloride reduced the yield of the final product. Premixing of acyl chloride with diisopropylethylamine is expected to yield a quaternary amide that while highly activated could be sterically hindered from reaction with **Z**(Act). Steric hindrance could also account for reduced yields from all coupling reagents other than acyl chlorides.

The assembled peptide was cyclized on the solid support with catalytic quantities of hydrochloric acid in 20% 2,2,2-trifluoroethanol/dichloromethane. Cyclization is complete within 15 min. Several events must occur for cyclization to occur including activation of the acetal group, removal of the acetone-protecting group, and hydrazone formation (Scheme 3). The rapid cyclization suggests that hydrochloric acid catalyzes the formation of an oxocarbonium ion that in turn reacts either with the imine nitrogen or the free amine. The acetone-protecting group is removed by hydrolysis. Since similar yields were obtained with either 1 or 5 equiv of hydrochloric acid, the poorly basic *N*-amino group is not significantly protonated under the reaction conditions. The methoxocarbonium ion may undergo an exchange reaction with 2,2,2-trifluoroethanol to give the more electrophilic, trifluroethoxocarbonium ion. However, 2,2,2-trifluoroethanol is not necessary for cyclization since the yields of **2** were only partially reduced when the cyclization reaction was carried out either in *N*-methylpyrrolidone or *N*,*N*-dimethylformamide.

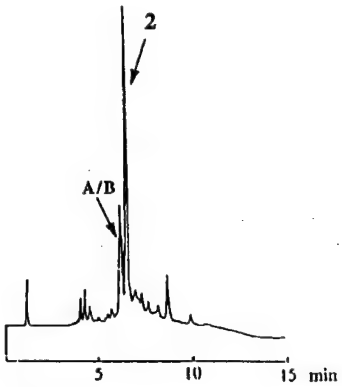


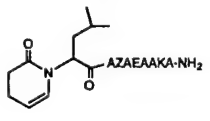
Figure 2. HPLC chromatogram of the crude product from the solid-phase synthesis of **2**. The major product is **2**. Two minor products, **A/B** are indicated. **A** (peak) is a dimer of **2**; **B** (shoulder) is an isomer of **2**.

Following cyclization, the peptide was cleaved from the Rink resin and protecting groups removed with trifluoroacetic acid. Analytical HPLC of the crude product (Figure 2) showed a single major product, **2**. Peptide **2** was purified by HPLC to give a final yield of 47%. Two minor products (**A/B**) were further purified and identified as a dimer (**A**) and an isomer of **2** (**B**) by mass spectroscopy. The dimer displays an NMR signal corresponding to that for a hydrazone proton while the isomer does not. Apparently, when **J** is activated, it reacts with **Z** on an adjacent chain. This is followed by cyclization to yield the "bis" compound. The isomer **B** is unidentified. However, a reaction between the oxocarbonium ion and the cL2 amide nitrogen would give a 6-membered heterocycle with a mass identical to **2**.[29]

[JLPZ]AEAAKA-NH_2 (**3**) was synthesized in the same manner as **2**. However, the final yield of **3** following two purification steps was lower (4.7%) than that for **2**; the major product was a dimer of **3**. This suggests that L-proline slows the cyclization rate of the monomer.

Since the biological activity of peptides is normally measured in aqueous media, and synthetic vaccines must withstand weeks of exposure to physiological conditions, the chemical stability of **2** was assessed under physiological conditions of temperature and pH by monitoring it in water at pH = 7.4, 37 °C by HPLC. After one month, no change was observed in the retention time, and there was little detectable degradation. No change was observed in **2** after 1 day in 10% D_2O/90% H_2O, pH 2.9,

(29) Proposed structure for **B**. Proposed by Barry Noar.



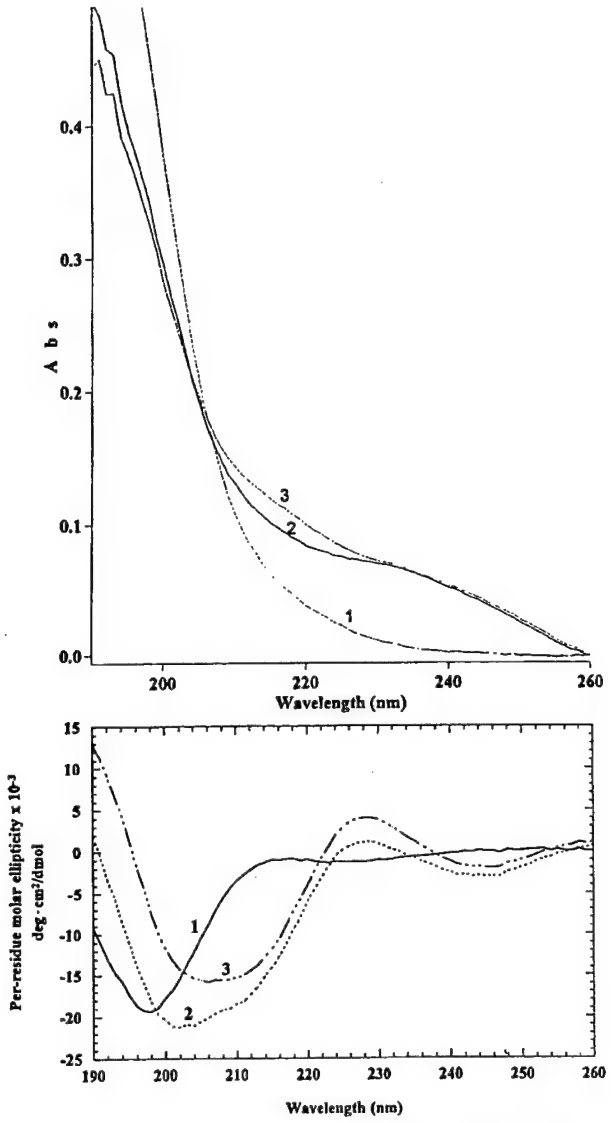


Figure 3. UV spectra for 10 μM **1**−**3** at pH 2.9, ambient temperature; path length is 1 cm (upper panel). CD spectra for 100 μM **1**−**3** at pH 2.9, 22 °C (lower panel).

conditions used for 2D NMR spectroscopy. However, **3** showed slow decomposition at pH 2.9 which became evident in NMR experiments. The lower yield and faster rate of decomposition of **3** may reflect ring strain. Peptides **2** and **3** showed little degradation under the acidic peptide cleavage conditions. However, the addition of ethanedithiol to the cleavage mixture rapidly destroys the hydrazone link and should be avoided in peptide deprotection reactions.

Conformational Analysis

CD Spectroscopy. Peptides **1**−**3** were analyzed by UV, circular dichroism (CD), and NMR spectroscopy. UV and CD spectra for **1**−**3** are compared in Figure 3. The CD spectrum for **1** in water at ambient temperature shows little evidence for helix while the CD spectra for **2** and **3** are complex. CD spectra are often analyzed in terms of UV spectra.[30] The difference in UV spectra for **1**−**3** indicates that the acylhydrazone link absorbs strongly in the UV with a broad band(s) centered at about 230 nm. The slightly stronger absorbance by **3** between 204 and 232 nm compared with **2** may reflect conformational effects on absorption by the conjugated acylhydrazone link.[31] The UV absorption band(s) attributable to the acylhydrazone are reflected in the CD spectra by positive mean residue ellipticity at 229 nm (**2**: $\Theta_{229} = 1064$) and 228 nm (**3**: $\Theta_{228} = 3990$) and negative mean residue ellipticity at 243 nm (**2**: $\Theta_{243} = -3005$) and 246 nm (**3**: $\Theta_{246} = -2030$).

The α-helix is characterized by a positive ellipticity at 190 nm and a double minimum at about 208 and 222 nm.[26,30] The CD spectra for **2** and **3** could include bands from both the hydrazone link and α-helix which would be difficult to separate since they would overlap. Also, the higher mean residue ellipticity for **3** compared with **2** at 228 and 246 nm suggests that the hydrazone link like the peptide bond contributes to CD spectra in a conformationally dependent manner that would be difficult to correct for. However, $[\Theta]_{190}$ which increases from −9161 (**1**) to 1857 (**2**) to 13003 (**3**) and a shifting minimum from 198 (**1**) to 202 (**2**) to 206 nm (**3**) are consistent with increasing α-helicity.

NMR Spectroscopy. Peptides **1**−**3** were principally analyzed by NMR spectroscopy. Several criteria have been developed for identifying α-helical peptides including upfield shifts in signals for main-chain amide NH protons, decreases in amide proton temperature coefficients and H−D exchange rates, decreases in $J_{\alpha N}$ coupling constants, and specific distance patterns in two-dimensional (2D) proton−proton nuclear Overhauser enhancement (NOE) spectra.[32,33] NOEs which appear as cross-peaks in 2D NOESY spectra provide the most important data since α-helices are characterized by sequential distances, $d_{NN}(i, i + 1)$, $d_{\alpha N}(i, i + 3)$, and $d_{\alpha\beta}(i, i + 3)$, which allow conformational analysis on a per residue basis.[33] NOE data are the least subject to mitigating factors that often complicate the interpretation of other spectral data and is considered the most reliable indicator of peptide structure in solution.[6,33,34]

NMR spectra were acquired on 1−20 mM peptides at pH 2.9, 22 °C. Although lower pH is expected to decrease the α-helical propensity of the peptide,[26] it optimally separates signals for the amide protons in NMR spectra which is important for the unambiguous assignments of NOEs. NMR studies on helical peptides are generally carried out at 5 °C [6] since the α-helix is significantly stabilized at lower temperatures.[26] However, low-temperature NMR requires specialized equipment that was not available to us. A comparison of chemical shifts and line shapes of the amide protons for 1 and 20 mM of **2** showed no discernible differences, ≤0.01 ppm, indicating that self-association is not significant at these concentrations. The resonance positions for all protons in **1**−**3** were assigned by using standard approaches to 1D ^{1}H NMR and 2D COSY, TOCSY, and ROESY (rotating-frame NOESY) spectra of undeuterated and deuterated analogues (see the Experimental Section and Supporting Information). Deuterated alanine analogues of the peptides (Table 2) were used to identify alanine amide proton signals for **1** and to isolate $d_{\alpha\beta}(i, i + 3)$ NOEs for **2**. The use of deuterated analogues of **2** had an additional advantage of reducing noise in the crowded regions of the ROESY spectra which enhanced signal/noise ratios for weak

(30) (a) Holzwarth, G.; Doty, P. *J. Am. Chem. Soc.* **1965**, *87*, 218−228. (b) Sears, D. W.; Beyechok, S. *Physical Principles and Techniques of Protein Chemistry;* Academic Press: New York, 1973; pp 460−466.

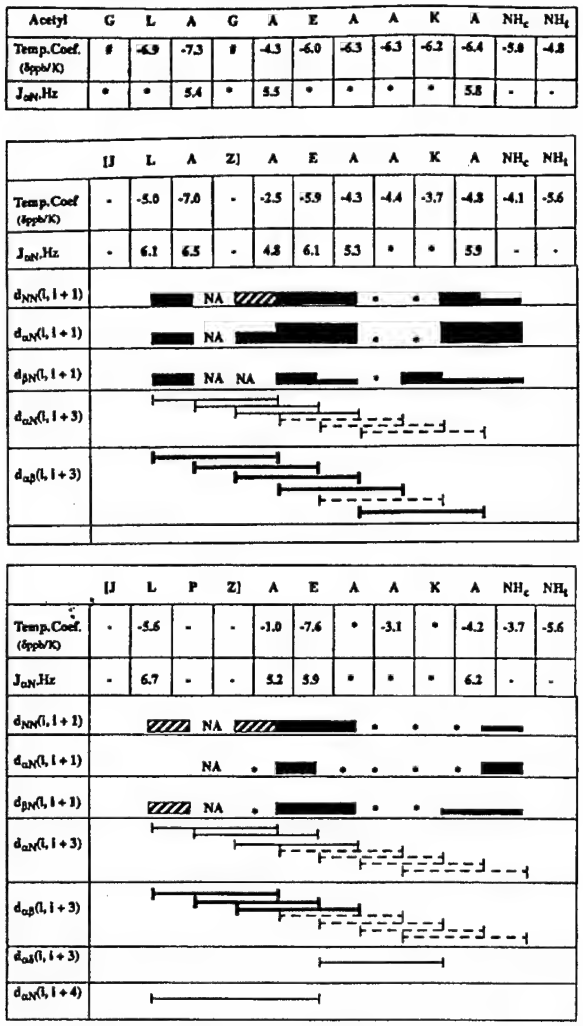
(31) Buchwald, M., Jencks, W. P. *Biochemistry* **1968**, *7*, 844−859.

(32) (a) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519−538. (b) Waltho, J. P.; Feher, V. A.; Merutka, G.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1993**, *32*, 6337−6347.

(33) Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715−740.

(34) Bradley, E. K.; Thomason, J. F.; Cohen, F. E.; Kosen, P. A.; Kuntz, I. D. *J. Mol. Biol.* **1990**, *215*, 607−622.

Table 3. Summary of NMR Data of **1**–**3**[a]



[a] Temperature coefficients were determined from linear regression analysis of five measurements between 21–45 °C with $R > 0.996$; the data is plotted in the Supporting Information. # indicate that the G1/G4 αN cross-peaks for **1** overlap in the TOCSY spectra; temperature coefficients for G1 and G4 are similar to those for E6. NH_c and NH_t are the carboxyl terminal cis and trans carboxamide protons. $J_{\alpha N}$ coupling constants are from 1D NMR spectra with a resolution of 0.33 Hz/datapoint. * and dashed lines indicate overlapping signals. Hatched bars are for NOEs from cL2 NH to cP3 δ CH_2, from cL2 β to cP3 δ CH_2 and from Z4 N=CH to A5 NH. NA, not applicable. Strong, medium, and weak NOEs are identified by the height of the filled bars; thick lines correspond to strong NOEs thin lines to weak NOEs.

NOEs. $J_{\alpha N}$ coupling constants and amide proton temperature coefficients for **1**–**3** and NOEs for **2** and **3** are summarized in Table 3.

Amide NH Region. Portions of 1D NMR spectra for **1**–**3** are compared in Figure 4. Signals for the main-chain amide protons of **2** and **3** show major differences in chemical shifts from corresponding protons for **1** with signals for the AAKA-NH_{cis} amide protons shifting upfield. The trend in signals for **2** are accentuated in **3** (Figure 5). The upfield shifts for the AAKA-NH_{cis} amide protons are predicted for hydrogen bonding that accompanies α-helix formation[35] with the larger differences for **3** indicative of greater helix stabilization.

The chemical shifts for the amide protons for **2** and **3** that are located near the acylhydrazone link shift in a discordant manner (Figure 5) which as discussed later may reflect shielding effects from the cyclic portion of the molecule. Differences in the signal for the hydrazone N=CH proton are also evident (Figure 4). Coupling constants for the hydrazone proton change from **2** (3.6, 8.2 Hz) to **3** (2.5, 9.5 Hz), indicating that the linker adjusts its conformation in response to the substitution of alanine (**2**) with proline (**3**).[36] These changes in conformation at the amino terminal end of **2** and **3** are reflected by changes at the carboxyl terminal end. NMR signals for the two C-terminal carboxamide protons (Figure 4) converge from a difference of 0.50 (**1**) to 0.42 (**2**) to 0.38 ppm (**3**). Both coupling constants and signal convergence provide sensitive and easily measurable parameters for conformational changes that must be occurring through the full length of **2** and **3**.

Hydrazone Link. Secondary amides often form mixtures of cis and trans isomers that interconvert slowly on the NMR time scale and are distinguishable in NMR spectra.[37] The data, however, provides strong evidence that the acylhydrazone (main-chain C–N link) adopts the trans configuration which relegates the hydrogen bond mimic (N–N link) to the required "*cis*" position (Figure 1). First, only one set of NMR signals is observed for **2** and **3**, indicating that only one geometric isomer is present (Figure 4). Second, no NOE was observed between the cA3 α or β protons and the **Z** methylene protons for **2**, whereas strong $d_{\alpha\alpha}(i, i + 1)$ NOEs are observed for cis peptide bonds.[38] Likewise, no NOE was found between the cP3 β protons and the **Z** α protons. It is possible that the hydrazone link could enforce a cis bond at another position, the most vulnerable being the cL2-cP3 peptide bond in **3**. The strong NOE between cL2 NH and cP3 δCH_2 for **3** (Figure 7), however, indicates that the cL2-cP3 peptide bond is trans.[39] There is also no evidence for cis isomerization along the main-chain peptide bonds in either the cyclic or linear portion of **2** or **3**.

Additional NOEs agree with the targeted α-helical conformation (Figure 1). Strong NOEs between the hydrazone N=CH proton and the **Z** methylene proton(s) (not shown) indicate that the vinyl N–N=CH proton is trans to the main-chain amide nitrogen. Also, an NOE between the N=CH proton and the A5 NH proton (Figure 6, **2** and **3**) and a medium strength cL2-cA3 $d_{\alpha N}(i, i + 1)$ NOE (Figure 7, **2b**) with similar intensity to the cL2-cA3 $d_{NN}(i, i + 1)$ NOE (Figure 6, **2**) support the structure.

Distances. Helices are characterized by sequential $d_{NN}(i, i + 1)$, $d_{\alpha N}(i, i + 3)$ and $d_{\alpha\beta}(i, i + 3)$ NOEs.[33] ROESY spectra for **2** and **3** show $d_{NN}(i, i + 1)$ NOEs, in every case where signal separation permits, for cL2-cA3, A5-E6-A7 and K9-A10-NH_{cis}, for the whole length of the peptide in accordance with helix stabilization (Figure 6). The strongest evidence for a helix is provided by the midrange $d_{\alpha N}(i, i + 3)$ and $d_{\alpha\beta}(i, i + 3)$ NOEs that span the length of **2** (Table 3). These NOEs span cL2-cA3-Z4 in the cyclic peptide and A5-E6-A7 in the appended peptide (Figures 7, 8a). Although $d_{\alpha\beta}(i, i + 3)$ NOEs for the cL2 and cA3 α protons are unambiguous, those for the α protons of the -ZAEAAKA- extension are potentially overlapped with other NOEs. To isolate these signals, alanine α and β protons were deuterated (**2a**) and a spectrum acquired to identify the remaining K9 NOEs (Figure 8b). Proton pairs from A5 (α

(35) Reymond, M. T.; Huo, S.; Duggan, B.; Wright, P. E.; Dyson, H. J. *Biochemistry* **1997**, *36*, 5234–5244.

(36) In a further test, glycine was substituted for L-leucine in **2** to give [JGAZ]AEAAKA-NH_2. A comparison of NMR data for [JGAZ]AEAAKA-NH_2 ($J_{N-CH-CH2}$ = 6.0 Hz, 6.0 Hz; C-terminal carboxamide protons, Δppm = 0.48) with that for **1**–**3** confirms the flexibilty of the linker. It indicates that glycine relaxes the conformation of the NucSite and reduces its ability to propagate a helix. Cabezas, E.; Satterthwait, A. C., unpublished data.

(37) Grathwohl, C.; Wüthrich, K. *Biopolymers* **1976**, *15*, 2025–2041.

(38) Chazin, W. J.; Wright, P. E. *J. Mol. Biol.* **1988**, *202*, 623–636.

(39) Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley: New York, 1986; pp 122–125.

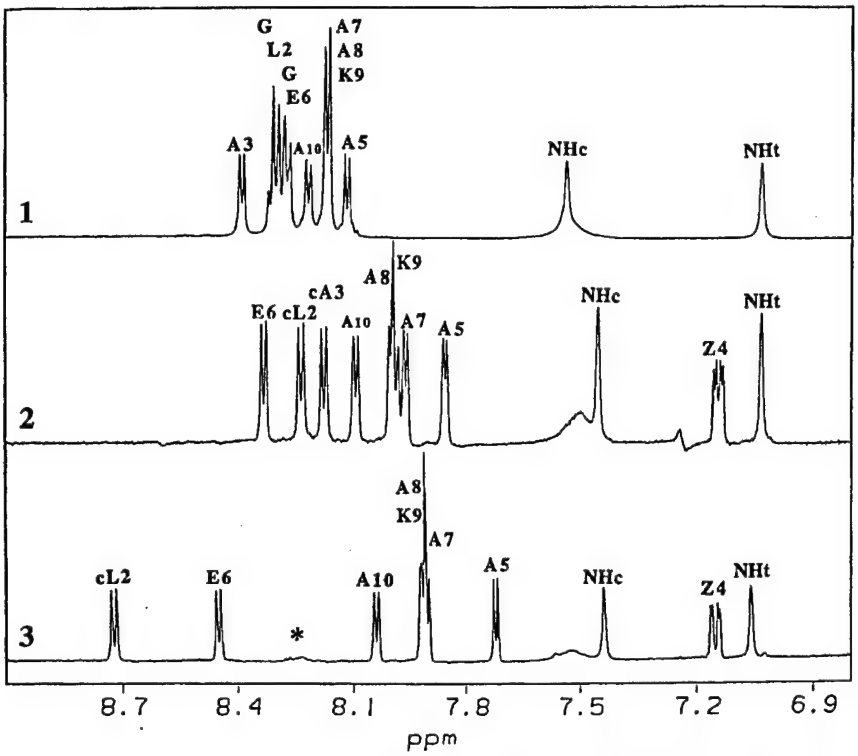


Figure 4. Amide proton region of 1D NMR spectra of **1**−**3** in 10% D_2O/H_2O at pH 2.9, 22 °C. Main-chain amide proton signals are identified for each amino acid according to the numbering scheme in Table 2. The signal for **Z**4 is from the N=CH proton. NH_c and NH_t are for the C-terminal cis and trans carboxamide protons.* Hydrolysis product that forms slowly at pH 2.9.

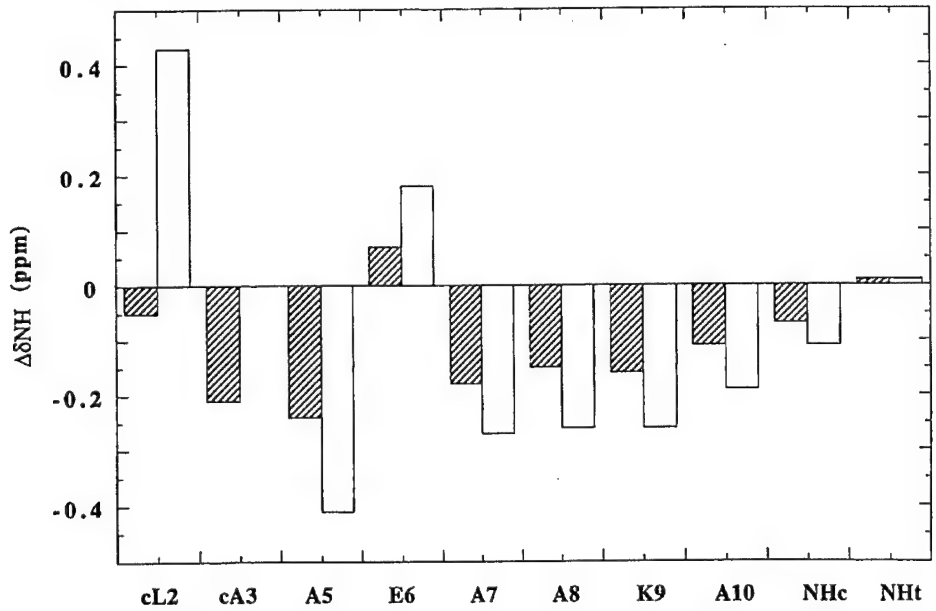


Figure 5. Differences in the chemical shifts for the main-chain amide protons of **2** (shaded bars) and **3** (open bars) compared with those for **1**. Data are from the spectra in Figure 4

proton) and A8 (β protons) and from A7 (α proton) and A10 (β protons) were then added back to give respectively **2b** and **2c** (Table 2) to unambiguously identify $d_{\alpha\beta}(i, i + 3)$ NOEs for these proton pairs (Figure 8c,d).

Midrange $d_{\alpha N}(i, i + 3)$ and $d_{\alpha\beta}(i, i + 3)$ NOEs that span cL2-cP3-Z4 and A5-E6-A7 were also observed for **3** (Figure 7,9) indicating that helix nucleation occurs in **3** in the same manner as in **2**. Overlapping signals prevented the assignment of midrange NOEs within the appended peptide of **3** except for a weak NOE that was observed between E6 αCH and K9 δCH_2. However, a similar pattern of upfield shifts in signals for the amide protons for AAKA-NH_{cis} indicates that **3** propagates a helix in the same way as **2** (Figure 5).

The 3_{10}-helix, α-helix and hybrid 3_{10}-α-helices are closely related structures that are difficult to distinguish by NMR spectroscopy.[20c,40] However, a comparison of distances for the

(40) Osterhout, J. J.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1989**, *28*, 7059−7064.

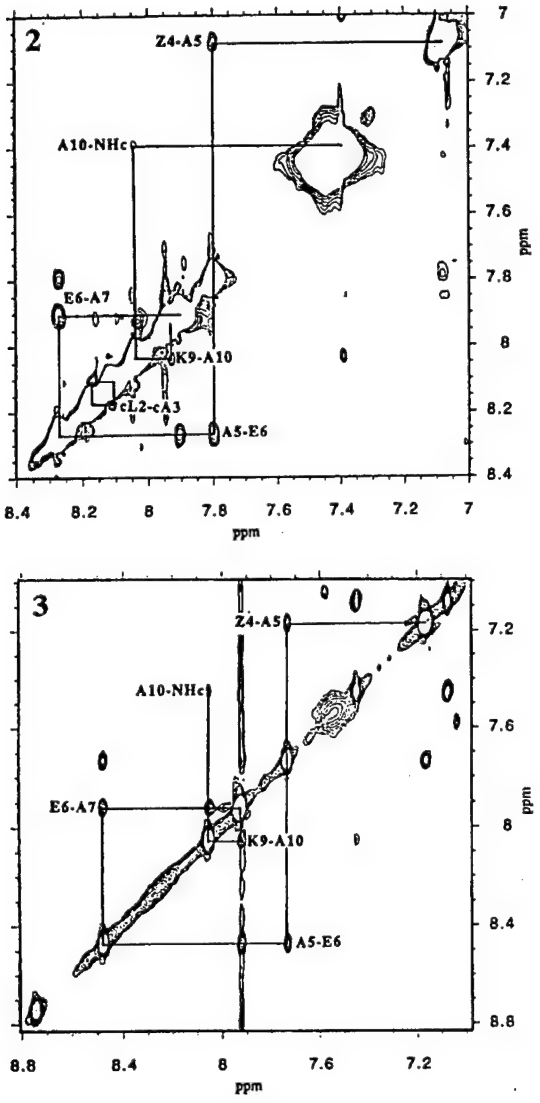


Figure 6. The amide proton region from a ROESY spectra of **2** in 10% D_2O/H_2O at pH 2.9, 19 °C (upper panel) and **3** at 22 °C (lower panel). Cross-peaks are for $d_{NN}(i, i + 1)$ NOEs identified according to the numbering scheme in Table 2. The Z4-A5 NOE is for $Z4^{N-CH}$-$A5^{N}$.

3_{10}- and α-helix lead to several predictions[20c,41] that can be tested with the NOE data for **2** and **3**. Two sets of distances for the 3_{10}- and α-helix, derived from similar models, were converted into relative NOE intensities (Table 4) for comparison. Since NOE intensities are to a first approximation proportional to $1/d^6$,[34] even small differences in models can affect the predicted NOE intensities. However, these uncertainties do not affect the following predictions.[41] (1) α-helices but not 3_{10}-helices are predicted to display very weak $d_{\alpha N}(i, i + 4)$ NOEs. (2) 3_{10}-helices are predicted to display very weak $d_{\alpha N}(i, i + 2)$ NOEs and weak $d_{\alpha N}(i, i + 3)$ NOEs. $d_{\alpha N}(i, i + 2)$ NOEs are predicted to be negligible for the α-helix. (3) For an α-helix, the relative intensities of NOEs are predicted to be $d_{\alpha\beta}(i, i + 3) \geq d_{NN}(i, i + 1) > d_{\alpha N}(i, i + 3)$ NOEs. For a 3_{10}-helix, the predicted relative intensities are $d_{NN}(i, i + 1) > d_{\alpha\beta}(i, i + 3) \approx 1.5 \times d_{\alpha N}(i, i + 3)$ NOEs.

The NOE data for **2** and **3** strongly support an α-helical structure with no indication of a 3_{10} structure. First, a very weak

(41) Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley: New York, 1986; p 162–166.

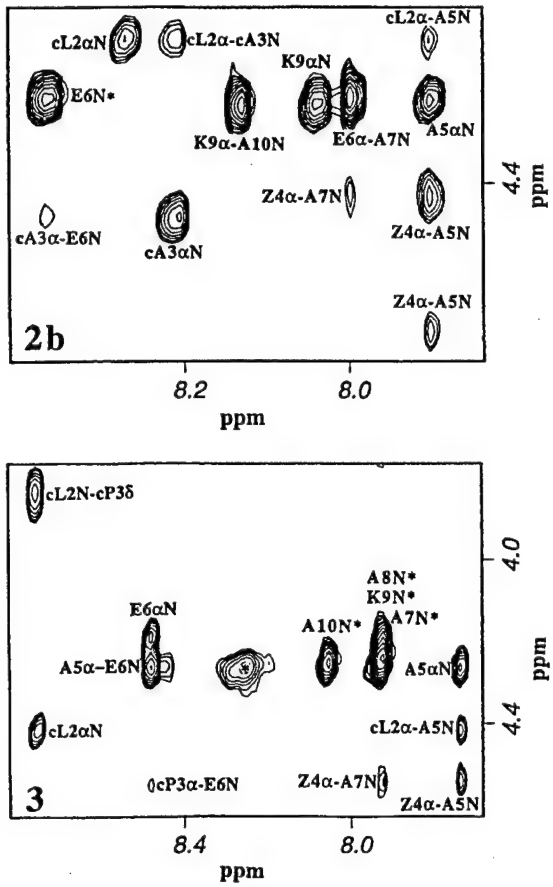


Figure 7. The fingerprint region from a ROESY spectra of **2b**, [JLAZ]-$A^3EA^4A^1KA^4$-NH_2 (upper panel) and **3** (lower panel) in 10% D_2O/H_2O at pH 2.9 at 22 °C. * indicates signals from a small amount of hydrolysis product of **3** that forms at low pH. N* indicates overlapping $d_{\alpha N}$ and $d_{\alpha N}$ $(i, i + 1)$ NOEs to the indicated NH.

$d_{\alpha N}(i, i + 4)$ NOE for cL2-E6 (**3**) is observed (see Supporting Information). Second, for **2** and **3**, $d_{\alpha\beta}(i, i + 3)$ NOEs for cL2-A5, cA3/cP3-E6, and Z4-A7 are stronger than $d_{NN}(i, i + 1)$ NOEs for A5-E6-A7 and far stronger than the corresponding $d_{\alpha N}(i, i + 3)$ NOEs (Figures 7–9, see Supporting Information for stack plot for **3**). Overlapping signals prevented a determination of whether $d_{\alpha N}(i, i + 2)$ NOEs were present or absent. Although $d_{NN}(i, i + 1)$ NOEs can potentially arise from unfolded forms and other conformers,[32b] any contributions from these sources would be additive. This would lower the observed ratio of $d_{\alpha\beta}(i, i + 3)/d_{NN}(i, i + 1)$ NOEs. Instead, $d_{\alpha\beta}(i, i + 3)$ NOEs are more intense than $d_{NN}(i, i + 1)$ NOEs, indicating that the α-helix is a major conformer. Since strong $d_{\alpha\beta}(i, i + 3)$ NOEs span [JL(A/P)Z] and -A5-E6-A7- of the appended peptide, the observed NOE ratios constitute strong evidence for α-helix nucleation with no apparent distortion from the acylhydrazone link. Similarly, strong $d_{\alpha\beta}(i, i + 3)$ NOEs for A5-A8 and A7-A10 in **2** indicate that the α-helix is propagated to the carboxyl terminal end.

In addition to the evidence for an α-helix, the observation of strong $d_{\alpha N}(i, i + 1)$ NOEs for -AEAAKA-NH_{cis} for **2** and **3** (Table 3, Figure 7) indicates that a fraction of the appended peptide is unfolded.[34] Strong $d_{\alpha N}(i, i + 1)$ NOEs are characteristic of β-extended forms (Table 4) and not predicted for an α-helix.[33] The NMR data for the constrained peptide is most readily interpreted in terms of a model that includes unfolded

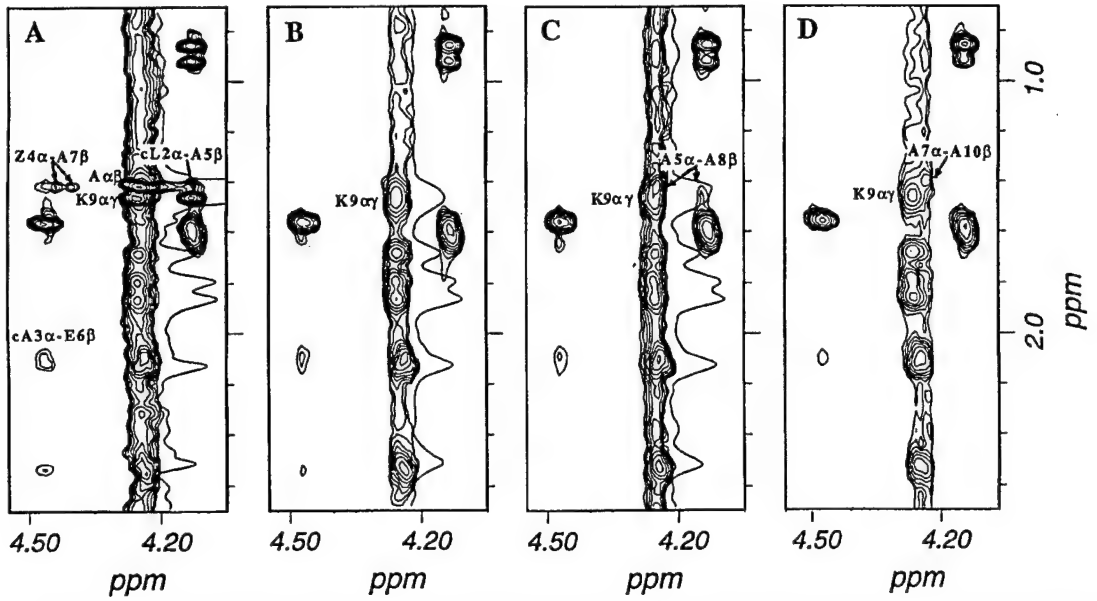


Figure 8. The $\alpha\beta$ region from ROESY spectra for **2** and deuterated analogues of **2** in 10% D_2O/H_2O at pH 2.9 at 22 °C. A 1D slice centered on the column for the K^{α} proton is superimposed on NOEs from the column to indicate the relative strengths of the NOE cross-peaks. A: **2**, identifies $cL2^{\alpha}$-$A5^{\beta}$, $cA3^{\alpha}$-$E6^{\beta}$, and $Z4^{\alpha}$-$A7^{\beta}$ NOEs. B: **2a**, [JLAZ]$A^4EA^4A^1KA^4$-NH_2, identifies the $K^{\alpha\gamma}$ NOEs by eliminating overlapping NOEs. C: **2b**, [JLAZ]$A^3EA^4A^1KA^4$-NH_2, identifies the $A5^{\alpha}$-$A8^{\beta}$ NOE. D: for **2c**, [JLAZ]$A^4EA^3A^4KA^1$-NH_2, identifies the $A7^{\alpha}$-$A10^{\beta}$ NOE.

Table 4. Distances in Å Between Designated Pair of Protons for Polyalanine in an Antiparallel β-Strand, 3_{10}-Helix and α-Helix[a]

distances, Å	β-strand[a]	3_{10}-helix[a]	3_{10}-helix[b]	α-helix[a]	α-helix[b]
$d_{NN}(i, i+1)$	4.3 (1.00)	2.6 (1.00)	2.59 (1.00)	2.8 (1.00)	2.57 (1.00)
$d_{\alpha N}(i, i+1)$	2.2 (55.00)	3.4 (0.20)	3.47 (0.17)	3.5 (0.26)	3.51 (0.15)
$d_{\alpha N}(i, i+2)$		3.8 (0.10)	3.79 (0.10)	4.4 (0.07)	4.32 (0.04)
$d_{\alpha N}(i, i+3)$		3.3 (0.24)	3.42 (0.24)	3.4 (0.31)	3.43 (0.17)
$d_{\alpha N}(i, i+4)$			5.54 (0.01)	4.2 (0.09)	4.18 (0.05)
$d_{\alpha\beta}(i, i+3)$		3.1 (0.35)	3.25 (0.26)	2.5 (1.97)	2.65 (0.83)

[a] Data is from Wüthrich et al.[33]; the $d_{\alpha\beta}(i, i+3)$ distance is for the closest approach of a β-methyl proton. [b] Based on helices generated by MSI Insight soft-ware. The alanine β-methyl protons are staggered relative to the alanine α-proton. The distance $d_{\alpha\beta}(i, i+3)$ is a weight-averaged distance to the three alanine β-methyl protons. Predicted intensities, in brackets, are relative to $d_{NN}(i, i+1)$ and based on $1/d^6$.

states in dynamic equilibria with folded states, including a full-length α-helix.

Dihedral Angles. Since $J_{\alpha N}$ coupling constants follow the Karplus equation, they are predicted to fall from >6 Hz for randomly coiled peptide (9 Hz for the extended chain) to about 4 Hz for an ideal α-helix.[42] The $J_{\alpha N}$ coupling constants for cL2 (6.1 Hz) and cA3 (6.5 Hz) in **2** are higher than predicted and indicate that the carbonyl oxygen atoms for J1 and cL2 tilt out about 20° on a weight-averaged basis from the axis of an ideal α-helix.[42] This tilt could reflect contributions from nonhelical forms of **2** which are not stabilized by hydrogen bonding. In **3**, proline ($\phi \approx -60°$)[15d] replaces alanine and shifts the equilibrium further toward the α-helix. The accommodation of a bulky L-proline with a constrained ϕ angle by **3** provides an important test for the targeted structure (Figure 1) that is well met. Proline's ability to enhance α-helix nucleation is probably due in part to better alignment of cL2 carbonyl oxygen along the helix axis.

The $J_{\alpha N}$ coupling constants for the appended peptide on **2** and **3** are not significantly reduced from those observed in **1**

(42) Calculated on the basis of the Karplus relationship that applies to ϕ angles in proteins; Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley: New York, 1986; pp 166–168.

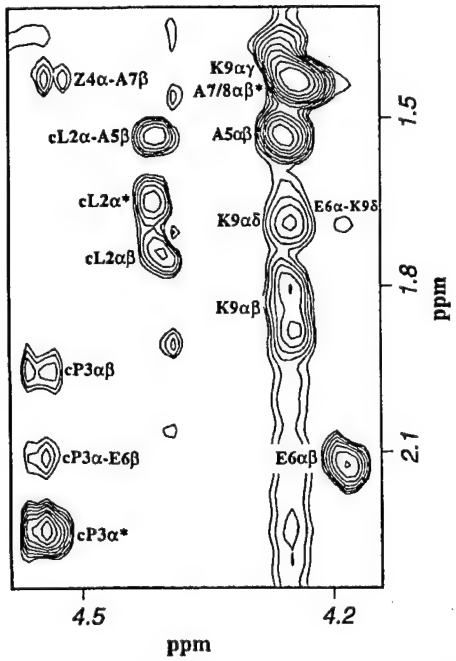


Figure 9. The $\alpha\beta$ region from a ROESY spectrum of **3** in 10% D_2O/H_2O at pH 2.9 at 22 °C. Identifies $cL2^{\alpha}$-$A5^{\beta}$, $cP3^{\alpha}$-$E6^{\beta}$, $Z4^{\alpha}$-$A7^{\beta}$ and a weak $E6^{\alpha}$-$K9^{\gamma}$ NOEs. α* indicates overlapping $\alpha\beta$ and $\alpha\gamma$ NOEs. $\alpha\beta$* indicates overlapping A7 and $A8^{\alpha\beta}$ protons.

(Table 3). However, coupling constants are weight-averaged values for mixtures of different conformers that could contribute in different ways to the observed values. It has been noted before that $J_{\alpha N}$ values are the least sensitive NMR parameter for detecting structure.[32]

Temperature Coefficients. α-Helices are characterized by sequential (i, $i+4$) hydrogen bonds that are associated with lowered H–D exchange rates, upfield shifts in signals for main-chain amide protons and decreases in the magnitude of amide proton temperature coefficients (<6 ppb/K).[32,35] Temperature

coefficients for **1**−**3** are summarized in Table 3; the data is plotted in the Supplementary Information.

The temperature coefficients for **1** are ≥ 6 ppb/K except for the A5 amide proton (4.3 ppb/K). This indicates that a fraction of the A5 amide proton is sequestered, probably by hydrogen bonding. The structure formed by **1** may be more local than propagated, although a small amount of helix cannot be ruled out.[43] The upfield shifts in signals for the amide protons of -AAKA-NH_{cis} for **2** and **3** (Figure 5) are accompanied by lowered temperature coefficients from approximately 6.3 ppb/K (**1**) to about 4 ppb/K (**2**, **3**) as expected of a peptide that has been stabilized as an α-helix. However, changes in chemical shifts and temperature coefficients for the preceding A5-E6 amide protons diverge from the norm (Figure 5). Specifically, the signal for the E6 amide proton moves downfield, and its temperature coefficient decreases slightly from 6.0 (**1**) to 5.9 (**2**) and then **increases** to 7.6 ppb/K (**3**). This would normally indicate that the E6 amide proton of **3** increases its exposure to water compared to **1**, which contradicts the conclusion from the NOE data that **3** nucleates α-helix formation. The corresponding data for A5 shows large upfield shifts in the amide proton signals (Figure 5) and a reduction in temperature coefficients to low values of 2.5 (**2**) and 1.0 ppb/K (**3**). This trend to divergence for the A5-E6 protons is accentuated further for a related peptide in 2,2,2-trifluoroethanol.[44] However, H−D exchange rates for A5-E6 amide protons were **both** depressed to identical low values in 2,2,2-trifluoroethanol.[44] Thus the large differences in temperature coefficients for the A5-E6 amide protons do not necessarily reflect differences in their exposure to solvent. Instead, it is probable that the chemical shifts for the amide protons of A5-E6 (**2**, **3**) are influenced by shielding effects from the hydrazone link. In an α-helical conformation the A5-E6 amide protons abut the cyclic peptide. This might account for downfield shifts in the signal for the E2 amide proton and enhanced upfield shifts in signals for the A5 amide proton. The higher temperature coefficients for E6 amide protons (**2**, **3**) could then arise from the added effect of a decrease in α-helix and shielding as temperature is raised (enhanced upfield shifts with increasing temperature) while an opposite effect is predicted for A5 amide protons (dampened upfield shifts).

The amide protons for -AAKA-NH_2 (**2** and **3**) which lie beyond the shielding effects of the cyclic peptide shift upfield as predicted for an α-helix. This agrees with the NOE data that indicates [JL(A/P)Z]AE nucleates α-helix formation. If the N termini of **2** or **3** were nonhelical, the observed upfield shifts for the -AAKA-NH_2 amide protons would not be expected. Data for the amide proton chemical shifts and temperature coefficients are thus consistent with increasing α-helix stabilization for **1**−**3**. Indeed, the divergence in chemical shifts and temperature coefficients for the A5-E6 amide protons which increases from **2** to **3**, more likely indicates the extent of α-helix nucleation.

(43) Percent α-helix is generally calculated from the mean residue ellipticity, Θ_{222}, and an extinction coefficient based on the length of the peptide (21). The reference state for 0% helix is $\Theta_{222} = 0$. Since $\Theta_{222} = -1268$ for **1** at pH 2.9, 22 °C, **1** forms 4% α-helix by the standard calculation: calculated from a ratio of $\Theta_{222}/\Theta_{max}$, where $\Theta_{max} = -39\,500 \times [1 - (2.57/n)]$ (= − 29 349 for $n = 10$).

(44) A comparison of chemical shifts, temperature coefficients and H−D exchange rates for the amide protons of acetyl-A5-E6(OEt)-E7(OEt)-E8-(OEt)-E9$(OEt)_2$ (**4**) was made with those for [JLAZ]-A5-E6(OEt)-E7(OEt)-E8(OEt)-E9$(OEt)_2$ (**5**) in 2,2,2-trifluroethanol and 2,2,2-trifluroethanol-d_1. The E6 amide proton of **5** shows a large disproportionate downfield shift relative to other amide protons when compared with **4**. The amide proton temperature coefficients for -AEEEE (**5**) are 0, 7.8, 4.0, 3.2, and 3.9 ppb/K respectively; Arrhenius, T.; Satterthwait, A. C., unpublished observations. The relative deuterium exchange rates for -AEEEE (**4**) are >10, >10, >10, >10, 10, respectively, while the corresponding rates for **5** are 1, 1, 2, 2, and 5, respectively.[16]

Table 5. Data is for **1**−**3** at pH 2.9, 22 °C[a]

parameters	peptide **1**	peptide **2**	peptide **3**
CD Spectral Data			
1. $[\Theta]_{190}$, deg·cm²/dmol	−9161	1857	13003
2. λ_{min}, nm	198	202	206
NMR Spectral Data			
1. δ NH upfield shift, Δppm			
A7 NH		0.18	0.27
NH_{cis}		0.07	0.11
2. $J_{N=CH-CH2}$, Hz		3.6, 8.2	2.5, 9.5
3. C-terminal carboxamide protons, Δppm	0.50	0.42	0.38
4. $R_{NN} = d_{NN}(i, i+1)/d_{\alpha N}(i, i+1)$			
for A5-E6			0.39
for E6-A7		0.23	
for A10-NH_{cis}			0.19
5. $R_{\alpha\beta} = d_{\alpha\beta}(i, i+3)/d_{\alpha N}(i, i+1)$			
for $cP3^{\alpha}$-$E6^{\beta}$/$A5^{\alpha}$-$E6^{N}$			1.15

[a] CD spectral data are from Figure 3. Differences in the chemical shifts for the amide protons for **2** and **3** are relative to **1**; data is from Figure 5. Coupling constants are from spectra in Figure 4. The differences in chemical shifts for the carboxamide protons were derived from temperature dependence data. The distance data are based on NOE volumes from spectra in Figures 6−9.

Correlates of α-Helicity. A comparison of the CD and NMR data for **1**−**3** reveals several measures that indicate or correlate with increasing α-helicity and provide measures for estimating the extent of α-helicity. These data which were noted above are compared in Table 5. Increasing $[\Theta]_{190}$ and λ_{min} in the CD spectra are characteristic of increasing α-helicity.[30] Upfield shifts in signals for the -AAKA-NH_{cis} amide protons are characteristic of α-helical peptides.[35] Coupling constants for the hydrazone N=CH proton diverge, whereas chemical shift differences between the two C-terminal carboxamide protons converge, reflecting conformational change over the full-lengths of **2** and **3**. The $d_{NN}(i, i+1)$ and $d_{\alpha\beta}(i, i+3)$ NOE intensities, representative of an α-helix, increase relative to $d_{\alpha N}(i, i+1)$ NOE intensities which are representative of unfolded forms.[32,34] The observation that $d_{\alpha\beta}(i, i+3)$ NOE intensities are similar to a $d_{\alpha N}(i, i+1)$ intensity (Table 5, stack plots for **3** in Supporting Information) despite the shorter $d_{\alpha N}(i, i+1)$ for extended forms (Table 4) suggests a very considerable fraction of α-helix. The general trend of the data indicates a progressive shift for **1**−**3** toward a larger fraction of α-helix. It indicates that the substitution of L-proline for L-alanine in **2** to give **3** improves α-helix stabilization by up to 2-fold. The decrease in upfield chemical shifts for the A10-NH_{cis} amide protons (Figure 5, Table 5) indicates that the α-helix frays at the end.

Estimates of the percentage of helix are generally based on CD measurements − the mean residue ellipticity at 222 nm, $[\Theta]_{222}$.[21,26] For **1**, $[\Theta]_{222} = -1268$. This could indicate a minor population of α-helix.[43] However, the determination of the percentage of helix from low ellipticity values is very dependent on a reference value for the unfolded or "coiled" state. It is not clear what this reference value should be for **1**. Different alanine-based peptides level off at different negative values of $[\Theta]_{222}$ as temperatures are raised, which suggests that the reference value for 0% helix is peptide-dependent.[26] In addition, the basic relationship between CD and helicity is not well understood.[45] A theoretical treatment by Manning et al.[45] suggests that a minimum length (2−3 turns) is required for an α-helix CD spectrum, that $[\Theta]_{222}$ is length-dependent for short peptides, and that intensities are significantly reduced by outward tilting of the carbonyl oxygen group from the helix axis. These factors could explain several differences between CD spectra and other

(45) . Manning M. C.; Illangasekare, M.; Woody, R. W. *Biophys. Chem.* **1988**, *31*, 77−86.

measurements.[6a,45] For **2** and **3**, additional problems arise, since the acylhydrazone link very likely contributes to $[\Theta]_{222}$ in a conformationally dependent manner that would be difficult to correct for.

The CD data leaves uncertain the status of **1** and the degree to which the hydrazone link stabilizes the peptide as an α-helix. The NMR data for **1**−**3**, however, identifies a new measure that is precise, easy to acquire, and very sensitive to changes in α-helical content. The decrease in Δppm for the two C-terminal carboxamide amide protons for **1**−**3** is striking (Table 5). Since **1**−**3** differ at their N termini, far removed in space from the C termini, the changes in the carboxamide protons are almost certainly related to the degree to which a full-length α-helix is formed. As the percentage of full-length helix increases from **1**−**3**, the difference in the chemical shifts for the carboxamide protons decreases from 0.50 to 0.38 ppm. Consequently, this simple measure can be used to detect small changes in the amount of full-length α-helix.

This prompted us to measure the differences in chemical shifts for the C-terminal carboxamide protons of acetyl-AEAAKA-NH_2. A small increase from Δppm = 0.50 ppm for **1** to Δppm = 0.52 ppm for acetyl-AEAAKA-NH_2 was observed for measurements made under the same conditions. This small difference, ΔΔppm = 0.02, indicates that acetyl-GLAG has very little effect on the C-terminal end of **1**. It provides a reference for measuring the effectiveness of [JLAZ]- and [JLPZ]- at propagating a full-length α-helix. The absolute change in ΔΔppm for the C-terminal carboxamide protons increases from **1** (0.02) to **2** (0.10) to **3** (0.14) indicating significant stabilization of the peptide as an α-helix in aqueous solution at 22 °C by the hydrogen bond mimic.

Discussion

The development of a solid-phase synthesis for replacing main-chain hydrogen bonds with the hydrazone link reduces the assembly time of α-helical peptides from 6 weeks to 3 days and improves overall yields at least 10-fold. The assembly process is also amenable to automation and multiple peptide syntheses. It requires little more effort to assemble an α-helical peptide than a peptide. Although yields depend on the individual peptide, the solid-phase synthesis makes accessible constrained peptides that were previously inaccessible. This development provides the wherewithal for assessing whether the hydrazone link is a hydrogen bond mimic and whether it can be used in a general sense to constrain peptides to a range of conformations defined by different hydrogen-bonding patterns (Table 1).

Hydrazone Link. Conformational analysis of **1**−**3** by NMR spectroscopy clearly indicates that the hydrazone link stabilizes a peptide as a full length α-helix in water. These studies further establish that the −N=CH− bond, like the hydrogen bond, positions itself "*cis*" to the main-chain peptide bond (Figure 1). NOE patterns for amino acids that span the hydrazone link (Table 3) indicate that it fits well into an α-helix in agreement with model building. A further indicator of the preference of the mimic for a "*cis*" position is provided by **3**. Whereas X-Pro peptide bonds form cis,trans equilibria,[37] the Leu-Pro peptide bond shifts to an all-trans geometry in **3** indicating that the hydrazone link satisfies the need for a cis bond better than L-proline. The hydrazone link also positions itself "*cis*" in larger loops,[17] demonstrating that this preference in **2** and **3** is not dictated by the small ring size. Presumably, this preference is steric; the −N=CH− link is less bulky than $-CH_2-CO-$.

An important property of the hydrogen bond is its unique flexibility. This provides the hydrogen bond with a capacity to adapt to different angles and lengths required by the different conformations that it defines in proteins.[14] The propylene extension of the hydrazone is also flexible and provides the linker with a capacity to adjust to different amino acids and peptide conformations.[17,36] While the −N=CH− bond length and angles are fixed and differ somewhat from the mean values for hydrogen bonds in proteins,[24] its preference for the "cis" position, small size, and the flexibility of the linker are sufficient for satisfying many of the properties required of a hydrogen bond mimic.

α-Helix Nucleation. The observation that a hydrazone link stabilizes an α-helix implies that enforcing one hydrogen bond is sufficient for stabilizing an α-helical conformation in water. The hydrazone link creates a nucleation site (NucSite) overcoming the most significant energy barrier to α-helix formation in agreement with the Zimm−Bragg model for α-helix formation.[46] Once the initial turn of an α-helix is made, further turns are propagated. The propagation of the hexapeptide extension, -AEAAKA-NH_2, as a full-length α-helix in water sets a lower limit on the number of turns (<2) that can be propagated as a full-length α-helix in water.

NucSites are functionally related to templates.[2,20,21] However, they differ in design. Indeed, template designs differ, reflecting a rich diversity of possibilities. NucSites are formally characterized by a main-chain link that can encompass different amino acids. Templates include side-chain linked di-[20a−e] and tri-prolines[20f] and nonpeptidomimetic helix inducers.[2,21] The type and extent of helix propagation is a function of the number of hydrogen bond acceptors on the nucleation site, their disposition in space, and the strength of hydrogen bonds between the appended peptide and the nucleation site versus the strength of hydrogen bonding to solvent.

NucSites and templates share a common feature, three carbonyl oxygen groups. Three hydrogen bond accepting groups can potentially propagate in addition to the α-helix, several other helices including the 3_{10}-, hybrid 3_{10}-α and mixed 3_{10}-α helices.[15d,47] The 3_{10}-helix which is often found at the carboxyl termini of α-helices in proteins is characterized by sequential ($i + 3 \rightarrow i$) hydrogen bonds and is more tightly wound.[15d,47a] One consequence of this is that side chains are projected in different orientations.[15d,47a] Although the α-helix is generally more stable,[47a] interconversion between the α-helix and 3_{10}-helix is readily achieved.[20c,47a] Since NucSites and templates present only three hydrogen bond accepting groups, a fourth hydrogen bond must be accommodated to nucleate an α-helical turn. This can be achieved in different ways.

Kemp and co-workers have reported extensive studies on the templated peptide, Ac-Hel_1-Ala_6-OH, in a variety of solvents.[20c,e] Ac-Hel_1 is characterized by three amide carbonyl oxygens. Back-bonding from the first three amino acids appended to the template could lead to a 3_{10}- or 3_{10}-α-helical turn.[20c] To propagate an α-helical turn, the amide linking the template to peptide must forego hydrogen bonding to the template in order for the next three amide groups to hydrogen bond to the three accepting carbonyl oxygen atoms on the template.[20c] NOE data for Ac-Hel_1 in water indicated helix stabilization.[20e] In 20 mol % trifluoroethanol-water, NOEs that were characteristic of both a 3_{10} and α-helix were detected,[20e] suggesting that solvent can participate in α-helix propagation.

(46) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526−535.
(47) (a) Toniolo, C.; Benedetti, E. *Trends Biochem. Sci.* **1991**, *16*, 350−353. (b) Toniolo, C.; Benedetti, E. *Macromolecules* **1991**, *24*, 4004−4009. (c) Karle, I. L.; Flippen-Anderson, J. L.; Gurunath, R.; Balaram, P. *Protein Sci.* **1994**, *3*, 1547−1555. Karle, I. L.; Gurunath, R.; Prasad, S.; Rao, R. B.; Balaram, P. *Int. J. Pept. Protein Res.* **1996**, *47*, 376−382 (d) Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J. *J. Am. Chem. Soc.* **1996**, *118*, 2744−2745

Bartlett and coworkers[21] devised another solution. They observed little helix propagation by N-linked templates, **9-N**. However, when the amide group that linked the peptide, -AAEALAKA-NH2, to the template was replaced with an ester, the resulting **S,S-9-O** templated peptide displayed an α-helix-like CD spectrum in water. They attributed the lack of helix propagation by the **9-N** template to an initial nonproductive 3_{10}-like hydrogen bond. By replacing the linking amide with an ester, the initial hydrogen bond was eliminated, making way presumably for the following amino acids to hydrogen bond to carbonyl oxygens on the template—a carboxylate anion, carbonyl oxygen and ester carbonyl oxygen were available for the propagation of an α-helix.

A different approach to template design was taken by Müller et al.[2] They described two templates that incorporate the initial amide link to the appended peptide into the template, either as a secondary lactam or imide which replaced the initial amide hydrogen. These templates provide, in addition, either two carbonyl oxygens and a "relay" lactam carbonyl oxygen or two carboxylate oxygens and a "relay" imide carbonyl oxygen which are available for α-helix propagation.

NucSites establish an α-helix by replacing the initial N-terminal (i, $i + 4$) connection with an acylhydrazone link. The two amide carbonyl oxygens and a third "relay" carbonyl oxygen atom from the acylhydrazone (Figure 1) provide hydrogen bonding sites for the following three appended amino acids, -AEA. Thus, each of the four amide protons in the first propagated turn of an α-helix are accommodated.

α-Helix Stability. The average size of an α-helix in a globular protein is about 11 residues[48a] with 88% of helices from a representative sample of proteins between 4 and 15 amino acids in length.[25a] Although a few native peptides in this size class form detectable levels of α-helices in water at low temperatures (3−7 °C),[22h,49] many, including longer peptides (>15 residues) from the helical regions of proteins do not.[50] NucSites, templates,[2,20,21] and other methods[22,23] might therefore find use in stabilizing native peptides as α-helices. However, the amino acid sequences of protein helices are diverse and include helix-destabilizing as well as helix-stabilizing amino acids. The helix propensities of amino acids have been reviewed.[48b,51] Model studies with alanine-rich peptides suggest that very little helix formation is expected for peptides not rich in alanine.[48b] On the other hand, sequence-dependent effects have been observed to play important roles in helix stability, suggesting a complicated reality.[40,52] To sort out these effects, it may be important to stabilize native peptides as helices in water. Consequently, it is of interest that a parallel study indicates that NucSites stabilize a native peptide sequence as an α-helix.[53] Stabilizing native peptides as α-helices could aid studies on protein folding by identifying determinants of helix stability[49,54] and enhance the activities of peptides for specific reactions[8d,22a,55] or for general peptide screening.[4,5b,c]

Wider Role. In principle, a hydrogen bond mimic could be used to stabilize peptides in different conformations defined by position and sequence (Table 1). The solid-phase synthesis for insertion of the hydrazone link into a peptide provides a facile method for exploring this possibility. Medium-sized loops have been prepared by extending **Z** with increasing numbers and varied sequences of amino acids before capping with **J** as is being reported separately.[9,17]

Conclusion

The hydrazone link was substituted for an ($i + 4 \rightarrow i$) hydrogen bond at the N terminus of a short peptide stabilizing it as a full-length α-helix in water at ambient temperature as indicated by NMR spectroscopy. The link is stable for long periods under physiological conditions of pH and temperature. The advantages of the current method are that (1) it provides the first solid-phase synthesis of an α-helix nucleation site, (2) the synthesis is rapid, convenient, and capable of high yields, (3) the nucleation site can accommodate different amino acids, and (4) the same synthetic protocol can be combined with multiple peptide synthesis procedures to synthesize different structures simultaneously.

This work establishes the hydrazone link as a hydrogen bond mimic, that a hydrogen bond mimic is sufficient for stabilizing a biologically relevant conformation in water, and that it is flexible enough to allow different amino acid sequences to exert their natural conformational preferences.

Experimental Section

General. Acetone, acetonitrile, dichloromethane, *N,N*-dimethylformamide, *N*-methylpyrrolidone, methanol, and 2-propanol from Burdick and Jackson were of high quality and were used without further treatment. Ethyl hydrazinoacetate hydrochloride, 1-hydroxybenzotriazole hydrate (HOBt), *N,N*-diisopropylethylamine, 2,2,2-trifluoroethanol, deuterium oxide, chloroform-d_1, and ninhydrin were from Aldrich. *N,N*-diisopropylethylamine was purified by distillation and stored over 4-Å sieves. Trifluoroacetic acid (TFA) was from Pierce Chemical Co. Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexaflurophosphate (PyBOP) and 9-fluorenylmethyl chloroformate (Fmoc-Cl) were from NovaBiochem. *N*-α-Fmoc-L-amino acids, *N,N*-dicyclohexylcarbodiimide (DIC), and 4-(2′,4′-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy resin (Rink resin) were from Advanced Chem Tech. Fmoc corresponds to the 9-fluorenylmethyloxycarbonyl group. L-Alanine-2-d_1, l-alanine-3,3,3-d_3 and L-alanine-d_4 were from MSD Isotopes. The pH test strips 0−14.0 (P4536) were from Sigma. All amino acids are L-amino acids. Proton chemical shifts are relative to tetramethylsilane except were noted.

Resin packets were made from style 67221030 polypropylene fabric from Synthetic Industries, Gainesville, GA, by sealing with a TISH-300 Impulse sealer from San Diego Bag and Supply. Shaking was carried out on an Eberbach 6010 two-speed reciprocating platform shaker.

5,5-Dimethoxy-1-oxopentanoic Acid Methyl Ester. The compound was synthesized according to Stevens and Lee.[27]

5,5-Dimethoxy-1-oxopentanoic Acid (J). The compound was prepared by adding 5,5-dimethoxy-1-oxopentanoic acid methyl ester (10 g, 56.8 mmol) to a solution of sodium hydroxide (3.028 g, 75.7

(48) (a) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomed. Struct.* **1992**, *21*, 95−118. (b) Chakrabartty, A.; Baldwin R. L. *Adv. Protein Chem.* **1995**, *46*, 141−176.

(49) (a) Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470−2474.

(50) Reymond, M. T.; Merutka, G.; Dyson, H. J.; Wright, P. E. *Protein Sci.* **1997**, *6*, 706−716.

(51) Woolfson, D. N.; Alber, T. *Protein Eng.* **1995**, *4*, 1596−1607. Schneider, J. P.; Lombardi, A.; DeGrado, W. F. *Folding Des.* **1998**, *3*, R29−R40.

(52) Lesk, A. M. *Nature* **1991**, *352*, 379−380; Serrano, J.; Sancho, J.; Fersht, A. R. *Nature* **1992**, *356*, 453−455. Zerkowski, J. A.; J. A.; Powers, E. T.; Kemp, D. S. *J. Am. Chem. Soc.* **1997**, *119*, 1153−1154. Williams, L.; Kather, K.; Kemp, D. S. *J. Am. Chem. Soc.* **1998**, *120*, 11033−11043.

(53) A peptide (underlined) from MSP-1, a *Plasmodium falciparum* merozoite protein, was inserted into acetyl-GLAGALFQKEKMAKA-NH_2 and [JL(A/P)Z]ALFQKEKMAKA-NH_2. 1D and 2D NMR experiments, similar to those reported for **1−3**, show that the constrained peptide is stabilized as an α-helix in 10% D_2O/H_2O at pH 5.0 at ambient temperature: Cabezas, E. and Satterthwait, A. C., unpublished data.

(54) Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−4255. Renold, P.; Kwok, Y. T.; Shimizu, L. S.; Kemp, D. S. *J. Am. Chem. Soc.* **1996**, *118*, 12234−12235.

(55) Chorev, M.; Roubini, E.; McKee, R. L.; Gibbons, S. W.; Goldman, M. E.; Caulfield, M. P.; Rosenblatt, M. *Biochemistry* **1991**, *30*, 5968−5974.

mmol) in 100 mL of an equivolume mixture of methanol in water and stirred for 2 h at room temperature. The reaction mixture was rotoevaporated, and the residue was dissolved in a mixture of 50 mL of water, added to 100 mL of ethyl acetate, and cooled to ice temperature. The mixture was vigorously stirred and acidified to pH 2−3 with 1 N hydrochloric acid. The pH was monitored by spotting pH 0−14 test strips. The organic and aqueous layers were immediately separated, and the aqueous layer was extracted with 100 mL of ice-cold ethyl acetate. The combined organic layers were washed with 100 mL of ice-cold brine, dried over magnesium sulfate, filtered, and rotoevaporated to yield 8.2 g (88%) of a colorless oil. ^{1}H NMR ($CDCl_3$) δ 1.55−1.8 (m, 4), 2.2−2.6 (m, 2), 3.33 (s, 6), 4.39 (t, 1). Anal. Calcd for $C_7H_{14}O_4$ C, 51.8; H, 8.7. Found C, 50.58, 50.65; H 8.88, 8.75. FAB MS Calcd (MNa^+) 185.0790; found 185.0793. **J** is stable for >1 year at − 20 °C.

(1-Methylethylidene-2-Fmoc)hydrazinoacetic Acid (Fmoc-Z(Act). The compound was synthesized by dissolving ethyl hydrazinoacetate hydrochloride (15.4 g, 100 mmol) in 300 mL of an equivolume mixture of water and acetone, refluxing for 10 min, and cooling to room temperature. Sodium hydroxide (8.4 g, 210 mmol) was added to the reaction mixture and stirred for 30 min at room temperature to hydrolyze the ester. Sodium carbonate (10.6 g, 100 mmol) was dissolved in 50 mL of water and added to the reaction mixture, and the mixture was cooled to ice temperature. A solution of 9-fluorenylmethyl chloroformate (25.9 g, 100 mmol) in 50 mL of dioxane was added dropwise to the reaction mixture during 1 h with stirring. Stirring was continued for 15 h while the reaction mixture came to room temperature. The reaction mixture was rotoevaporated and the residue suspended in 200 mL of ice-cold ethyl acetate. This mixture was then vigorously stirred and acidified to pH 3 by adding 1 N HCl. The two phases were separated, and the aqueous phase was extracted with 200 mL of ethyl acetate. The ethyl acetate extracts were combined, extracted with brine, and dried with magnesium sulfate. After filtration and rotoevaporation to dryness, the product was dissolved in 100 mL of acetone, refluxed for 10 min, and again rotoevaporated to dryness. Crystallization from ethyl acetate with hexane yielded 31.4 g (80.7%) of white crystals: mp 146−148 °C; ^{1}H NMR (acetone-d_6) δ 1.5−1.75 (br s, 2), 1.7−2.0 (br s, 4), 4.19 (m, 2), 4.27 (m, 1), 7.33 (dd, 2, J = 7.4, 7.4), 7.42 (dd, 2, J = 7.4, 7.5), 7.68 (d, 2, J = 7.4), 7.87 (d, 2, J = 7.6), 11.06 (s, 1). Anal. Calcd for $C_{20}H_{20}N_2O_4$ C, 68.17; H, 5.72; N, 7.95. Found C, 67.97; H, 5.71; N, 7.85. FAB MS Calcd (MH^+) 353.1422; found 353.1511. Fmoc-Z(Act) is stable when stored dry at 4 °C.

***N*-α-Fmoc-L-alanine-2-d_1, *N*-α-Fmoc-L-alanine-3,3,3-d_3 and *N*-α-Fmoc-L-alanine-d_4.** The compounds were prepared according to Lapatsanis et al.;[56] mp 142−143 °C [*N*-α-Fmoc-L-alanine, lit. = 142−143 °C].

***N*-α-Fmoc-L-alanine chloride and *N*-α-Fmoc-L-proline chloride.** The compounds were prepared by the procedure of Carpino et al.[28] The crystalline products were dried under vacuum to remove trace amounts of hydrochloric acid. Both compounds were >97% pure by the methanol test.[28]

Fmoc-Z(Act)-Ala-Glu(*t*-But)-Ala-Ala-Lys(*t*-Boc)-Ala-Rink Resin. The compound was assembled on 4-(2′,4′-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy resin (Rink resin) (100 mg, 0.046 mmol) sealed in a 2 cm^2 polypropylene mesh packet.[57] The packet was included among several packets for simultaneous multiple peptide synthesis.[57] The resin packets were placed in capped 30-mL or 60-mL wide-mouth polyethylene bottles for common wash and deprotection steps. Wash solutions were added to the packets in the bottle, shaken for 1 min on a platform shaker, and poured off. Single packets were placed in capped 2-mL micro tubes for coupling reactions. Otherwise, when several packets required the same amino acid, they were combined and reaction solutions scaled up for coupling in a capped bottle. The following protocol reports the reagents required for a synthesis with one packet.

A standard coupling cycle involved the following steps: (1) The resin was swelled in its packet by shaking with 5 mL of dichloromethane and then washed with 5 mL of *N*,*N*-dimethylformamide. (2) The Fmoc-protecting group was cleaved from the protected peptide−resin with 5 mL of 20% piperidine in *N*,*N*-dimethylformamide for 10 min, and the resin was washed four times with 5 mL of *N*,*N*-dimethylformamide, once with 5 mL of 2-propanol, once with 5 mL of dichloromethane, and once with 5 mL of *N*-methylpyrrolidone. (3) Coupling was carried out with *N*-α-Fmoc-L-amino acid (0.092 mmol) or Fmoc-**Z**(Act) (32.5 mg, 0.092 mmol), 1-hydroxybenzotriazole (12.4 mg, 0.092 mmol), benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexaflurophosphate (47.9 mg, 0.092 mmol), and *N*,*N*-diisopropylethylamine (0.032 mL, 0.184 mmol) in 1 mL of *N*-methylpyrrolidone by shaking for 30 min. After the coupling reaction, the resin was washed twice with methylene chloride and air-dried. A small amount of resin was removed from the resin packet and tested for primary amine by the Kaiser ninhydrin test.[58] The coupling step was repeated if necessary.

J-Leu-Ala-Z(Act)-Ala-Glu(*t*-But)-Ala-Ala-Lys(*t*-Boc)-Ala-Rink Resin. Peptide assembly was continued from above in exactly the same manner except for the coupling (step 3) of Fmoc-L-alanine chloride to Z(Act). Since the acetone-protecting group on Z(Act) is labile to strong acids, it is critical that Fmoc-L-alanine chloride be freed of adventitious hydrochloric acid by briefly drying it under vacuum. *N*-α-Fmoc-L-alanine chloride (30.3 mg, 0.092 mmol) was dissolved in 1 mL of *N*-methylpyrrolidinone, added to the resin packet, and briefly shaken before adding 0.016 mL of *N*,*N*-diisopropylethylamine (0.092 mmol). Shaking was continued for 15 min. The Kaiser ninhydrin test is not applicable for this reaction. Fmoc-L-leucine (32.5 mg, 0.092 mmol) and then **J** (17 mg, 0.092 mmol) were coupled to the peptide resin using the coupling protocol for Fmoc-amino acids. Following the addition of **J**, the resin packet was washed four times with 5 mL of methylene chloride and air-dried.

[J-Leu-Ala-Z]-Ala-Glu(*t*-But)-Ala-Ala-Lys(*t*-Boc)-Ala-Rink Resin. The dry resin packet from the previous step was shaken for 15 min in a 15 mL solution of 20% 2,2,2-trifluoroethanol in methylene chloride that had been premixed with 0.02 mL 4 N hydrochloric acid in dioxane (5.3 mM final). The resin packet was washed four times with 5 mL of dichloromethane. A slight pink coloring of the resin was used as an indicator that sufficient acid had been added for cyclization.

[J-Leu-Ala-Z]-Ala-Glu-Ala-Ala-Lys-Ala-NH$_2$ (2). The cyclic peptide-resin from the previous step was added to 15 mL of 10% trifluoroacetic acid in dichloromethane in a capped polypropylene bottle and shaken for 15 min. The solution containing the cleaved protected peptide was removed and the packet treated again in the same manner. The acidic solutions were combined and rotoevaporated to leave a residue that was treated with 2 mL of trifluoroacetic acid for 15 min. Peptide was precipitated from the acid solution by adding 30 mL of cold diethyl ether. The precipitate was centrifuged, washed with cold diethyl ether and dissolved in 1 mL 10% acetonitrile in water.

The crude product was initially analyzed by high-pressure liquid chromatography (HPLC) on a Merck Lichrosorb analytical C-18 column (4.6 × 25 mm, 10 micron). Products were eluted with a 100% water−100% acetonitrile gradient in 0.1% TFA over 15 min at 2 mL/min. The chromatogram (Figure 2) showed one major product, **2**, and two minor products, **A** and **B**. These products were separated by HPLC on a Vydac preparative C-18 column (201TP1022, 2.2 × 25 cm, 10 μm) by eluting with a 100% water−60% acetonitrile gradient in 0.1% trifluoroacetic acid over 40 min at 5 mL/min. Fractions were lyophilized to yield 19.2 mg of **2** (47% yield, based on initial resin loading). **2** was identified by NMR spectroscopy and FAB-MS: Calcd (MH^+) 896; Obsd 896. HR FAB-MS: Calcd (M^+Cs^+) 1027.3977; Obsd 1027.4019

The **A** and **B** fractions from several HPLC runs were combined and purified further for analysis. **A** was identified as a dimer of **2** by low resolution (LR) FAB MS: Calcd (MH^+) 1791; Found 1791. **B** was identified as an isomer of **2** by FAB MS: 895.5001, Found 895.5109. ^{1}H NMR (10% D_2O/90% H_2O) of **A** but not **B** showed a signal at 7.21 ppm (dd, 1) that is characteristic of a hydrazone N=CH proton.

[J-Leu-Pro-Z]-Ala-Glu-Ala-Ala-Lys-Ala-NH$_2$ (3). The compound was synthesized by the same procedure as **2** but with *N*-α-Fmoc-L-proline chloride used in place of *N*-α-Fmoc-L-alanine chloride. Rink's resin (1 gm, 0.42 mmol) was sealed in six resin packets for this

(56) Lapatsanis, L.; Milias, G.; Froussios, K.; Kolovos, M. *Synthesis* **1983**, 671−673.

(57) Houghten, R. A.; De Graw, S. T.; Bray, M. K.; Hoffman, S. R.; Frizzell, N. D. *Biotechniques* **1986**, *4*, 522−528.

(58) As described in Stewart, J. M.; Young, J. D. *Solid-Phase Peptide Synthesis*, 2nd ed.; Pierce Chemical Co.: Rockford, IL, 1984; p 105.

synthesis. The coupling reactions were scaled up proportionately. The wash and reaction steps were carried out in 60-mL Nalgene wide-mouth polyethylene bottles. Cyclization was carried out by shaking the six resin packets for 15 min in 40 mL of 20% 2,2,2-trifluoroethanol in dichloromethane that had been premixed with 0.105 mL 4 N hydrochloric acid in dioxane. The product was cleaved from the resin by scaling up the procedure used for **2** by 10-fold.

Two major products were separated by HPLC with multiple runs on a Merck Lichrosorb semipreparative C18 column (10 × 250 mm, 7 μm) with a 100% water−100% acetonitrile gradient in 0.1% trifluoroacetic acid over 30 min at 3 mL/min. A dimer of **3** eluted at 19 min, while **3** eluted at 19.7 min. The fractions of **3** were combined, lyophilized, and purified by repeating the HPLC runs to yield 16.5 mg (4.3%); HR FAB-MS: Calcd (MH^+) 921.5158; Obsd 921.5172. The dimer of **3** was obtained in 23% yield from the first purification; LR FAB MS: Calcd (MH^+) 1842, Obsd 1842.

Acetyl-GLAGAEAAKA-amide (1). The compound was synthesized in a resin packet by using the coupling cycle and cleavage procedures described above. The peptide was acetylated on the resin by first removing the final Fmoc-protecting group with 20% piperidine in *N,N*-dimethylformamide and then washing and shaking the peptide-resin with 15% acetic anhydride in *N,N*-dimethylformamide for 30 min. The peptide was purified by HPLC as a single peak and identified by NMR spectroscopy and LR FAB MS: Calcd (MH^+) 899; Obsd 899.

Acetyl-AEAAKA-amide. The compound was synthesized and purified in the same manner as **1**.

Deuterated analogues, 1a−d and 2a−c. The compounds were synthesized in resin packets as described above. Resin packets were combined for common steps and separated for individual coupling reactions[57] with *N*-α-Fmoc-L-alanine-2-d_1, *N*-α-Fmoc-L-alanine-3,3,3-d_3, and *N*-α-Fmoc-L-alanine-d_4. Products were purified by HPLC and their compositions confirmed by comparing their 1D NMR spectra with the 1D NMR spectra for the fully characterized, undeuterated compounds.

Spectroscopy. Purified peptide samples were lyophilized at least twice to remove trace amounts of TFA, weighed, and dissolved in doubly distilled water or 90% H_2O/10% D_2O. The aqueous samples were adjusted to pH = 2.9 by adding small quantities of HCl and NaOH. UV spectra were recorded in a 1-cm quartz cell (1-cm path length) with a Varian Cary 3Bio UV/Vis spectrometer. CD spectra were recorded on samples in a 0.2-cm quartz cell (0.2-cm path length) with an Aviv model 62ADS circular dichroism spectrometer. The observed spectra were corrected by subtracting a spectrum for water. Molar ellipticities for **1−3** are calculated for 10 residues.

Most NMR spectra of peptides were acquired with a Bruker AMX-500 spectrometer. Except where noted, all spectra were recorded on approximately 20 mM peptide, pH 2.9 at a probe temperature of 22 °C. Chemical shifts are relative to internal dioxane at δ = 3.75 ppm.

1D NMR spectra in 90% H_2O/10% D_2O were run with presaturation of water. $J_{\alpha N}$ coupling constants for amide protons were determined at 0.34 Hz/data point after zero filling. 2D ROESY experiments were run in the phase-sensitive mode using TPPI for quadrature detection in f_1. In a typical experiment, 48 transients of 1024 data points were acquired in the f_2 dimension at a spectral width of 5555 Hz and 512 t_1 increments which were zero filled to 1024 data points.

2D ROESY spectra were acquired for **2** with mixing times of 100, 200, and 400 ms and compared with a 2D NOESY spectrum on the same sample. NOEs were far more intense in the ROESY spectra than in the NOESY spectrum, indicating that the product of the Larmor precession frequency and the correlation time ($\omega\tau_c$) is close to unity. Strong NOEs with high signal/noise ratios were proportional to the mixing time. The NOE data that is reported is from 2D ROESY experiments carried out with a mixing time of 400 ms.

Temperature coefficients were determined from 1D and 2D TOCSY experiments run on a Bruker AM-300 spectrometer. Spectra were acquired from samples in 90% H_2O/10% D_2O which was suppressed using a jump-and-return selective excitation sequence. Probe temperatures were determined by using a methanol standard.

1D and 2D data was processed and analyzed on a Silicon Graphics Iris Indigo XS24 computer using Molecular Simulation, Inc. (MSI) Felix 95 software.

Acknowledgment. We acknowledge Dr. Lin-chang Chiang for his initial synthesis of Fmoc-**Z**(Act). We thank Drs. Thomas Arrhenius and James Tam for helpful conversations. Dr. John Waltho aided us in obtaining NMR measurements. The manuscript benefited from reviews by Drs. Elena Dovalsantos and Philip Dawson and from thoughtful suggestions by the final reviewers. This work was supported by grants from the United States Agency for International Development (DPE-5979-A-00-1035-00) and the National Institutes of Health (AI19499). This is publication 11533-MB.

Supporting Information Available: Tables of ^{1}NMR chemical shift data for **1−3**, the amide proton region from 1D NMR spectra for **1a−d**, fingerprint region from a 2D TOCSY spectrum of **1**, 2D TOCSY spectrum of **2**, stack plots and a 1D slice from a 2D ROESY spectrum of **3** and plotted data for the determination of temperature coefficients for **1−3** (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA983212T